

# SOME ASPECTS OF HINDI PHONOLOGY

A Thesis Submitted  
In Partial Fulfilment of the Requirements  
for the Degree of  
DOCTOR OF PHILOSOPHY

By  
ARUN CHANDRA PANDEY

*to the*

DEPARTMENT OF HUMANITIES AND SOCIAL SCIENCES  
INDIAN INSTITUTE OF TECHNOLOGY KANPUR  
DECEMBER, 1979

To the memory of my mother,

Smt. Champawati Pandey

This dissertation is humbly dedicated.

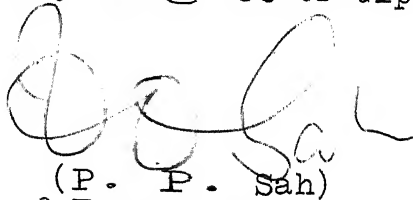
HSS-1978-D-PAN-SOM

I.I.T. KANPUR  
CENTRAL LIBRARY  
Acc. No. A 63064

12 AUG 1980

## CERTIFICATE

This is to certify that the thesis "Some Aspects of Hindi Phonology" submitted by Shri Arun Chandra Pandey in partial fulfilment for the degree of Doctor of Philosophy to the Indian Institute of Technology, Kanpur, is a record of bonafide research work carried out by him under my supervision and guidance. The results embodied in this thesis have not been submitted to any other university or institute for the award of any degree or diploma.



(P. P. Sah)  
Professor of English (Linguistics)  
Department of Humanities  
and Social Sciences  
Indian Institute of Technology  
Kanpur



## CERTIFICATE

This is to certify that Mr. Arun Chandra Pandey has completed all the course requirements for the Ph.D. programme in English (Linguistics).

He completed the following courses at I.I.T. Kanpur:

H-Eng 710	Approaches to the Study of Literature
H-Eng 715	Methodology of Teaching and Research in Literature
H-Phy 784	Applied Social Psychology

He was given advance standing in the Ph.D. programme as a result of his having completed the following courses at the University of California at Los Angeles and the Michigan State University:

Lin-100	Introduction to Linguistics (UCLA)
TESOL	Teaching of English as a Second Language (UCLA)
Lin-831	Grammatical Analysis (MSU)
Lin-321	Phenological Analysis (MSU)
Lin-403	Morphology (MSU)
Lin-880	Linguistics Seminar (MSU)
Lin-835	Semantics of Natural Language (MSU)
	Introduction to Historical Linguistics (UCLA)
LOA 490	Special Projects (Marathi) (MSU)
LOA 299	Special Projects (Marathi) (MSU)
Lin 860	Special Projects (MSU)

Mr. Arun Chandra Pandey was admitted to the candidacy of the Ph.D. degree in January 1977 after he successfully completed the written and oral qualifying examinations.



Head

Department of Humanities  
and Social Sciences  
I.I.T. Kanpur



Convener

Departmental Post-graduate Committee  
Department of Humanities  
and Social Sciences  
I.I.T. Kanpur

## ACKNOWLEDGEMENTS

This dissertation would not have been possible but for my friend, Sri Satya Kam Tandon. It was he who by his constant goading and arguing forced me to take the irrevocable step of starting this work. In fact, I dictated the first draft to him.

To thank my guide, Professor P. P. Sah, I do not have enough or appropriate words. My dissertation has throughout caused more trouble for others than for me, and being in the unenviable position of my guide, he had more than his share of them. He has most thoroughly and painstakingly gone through all of my drafts, noting down his comments and suggestions, discussing them with me, checking my wayward impulses, and gently and persuasively bringing me back to the right track. I have immensely benefited from his remarks and much of the tightening up of the argumentation is due to him. As I have not been able to match him in either thoroughness or perspicacity, the dissertation still suffers from a number of shortcomings, both regarding the content and the style.

Thanks are also due to my teachers from whom I learned my linguistics. In this connection I would like to make a special mention of Dr. Raimo Anttila of UCLA, Dr. David Stampe of Ohio State University, Columbus, Dr. John Ritter,

Professor Julia Falk and Dr. V. S. Khokle of Michigan State University, East Lansing, where I did my graduate work in linguistics. My interest in Indo Aryan is entirely due to Dr. Khokle, who was always ready to discuss questions regarding Sanskrit and Indo Aryan linguistics with me. Needless to say none of them is responsible for my imperfect understanding of the subject.

I should also like to thank my fellow students in the graduate program at MSU from whom I have learned a lot through discussions both in and out of class. Thanks are also due to Sarah Ulrey, whom I taught Hindi at MSU, for making me aware of medial consonant gemination in Hindi, a topic which is included in the present work.

I am greatly indebted to Professor K. N. Sharma, the Head of my Department at IIT Kanpur, who provided me with the maximum of free time for completing this work.

Thanks are due to Mr. Somnath Biswas and Dr. A. Sircar for helping me with Bengali data. I am also greatly indebted to Dr. B. N. Patnaik for information regarding Oriya.

I would like to acknowledge here the debt that I owe to my wife, Rajni, for the constant encouragement, and all possible help, that she has given me throughout the period this work was in progress. In fact, the tension has been much greater for her than for me.

Thanks are also due to Sri T. Dhanpal for his help in the onerous task of proof-reading and to Sri T. Chacko and Km. Kiran Dhingra for general encouragement that they have given me. I should like to thank Drs. P. N. Shukla and Prem Sagar also for their help.

Thanks are also due to Sri Nihal Ahmad for typing the manuscript, Sri Sudama Prasad for cyclostyling, and Sri Ram Adhar Rai for his help in preparing the copies.

## CONTENTS

	Page
Synopsis	viii-xii
Chapter 1 Introduction	1-15
Chapter 2 On Aspiration and the Nature of <u>h</u>	16-56
Chapter 3 On Glides and Diphthongs	57-67
Chapter 4 On Vowel Quantity and Vowel Nasalization	68-120
Chapter 5 On Mid Vowels	121-165
Chapter 6 On Word Final <u>a</u>	166-189
Chapter 7 On Syllable Structure	190-205
Chapter 8 Conclusion	206-220
Appendix	221-225
Bibliography	226-234

## SYNOPSIS

Name : ARUN CHANDRA PANDEY

Degree : Ph.D.

Institution : Indian Institute of Technology, Kanpur

Month and Year : December, 1979

Title : Some Aspects of Hindi Phonology

In this dissertation we have examined some aspects of Hindi phonology from the generative point of view. The topics chosen are of considerable interest not only for Hindi but also for Modern Indo Aryan languages (NIA) in general. The dissertation has eight chapters and one appendix.

Chapter One, the Introduction, sets out the theoretical orientation of the study. Although we follow the SPE framework in the main, there are certain deviations in the matter of the set of distinctive features used. In general, we have tried to avoid abstract solutions in favour of concrete ones, all other things being equal. We have also argued for bringing in historical information as well as facts from dialects and related languages in our analysis.

Chapter Two deals with the nature of h in Hindi. In the first section we take up the question of aspiration and examine both the unit segment and the cluster approach. We try to show that the cluster approach to aspiration (i.e.,

deriving  $C^h$  from the sequence  $C + \underline{h}$ ) is far superior to the unit segment approach in that it helps in understanding both the synchronic and the diachronic aspects of aspiration in Hindi in a much better way and also explains some aspects of Sanskrit phonology. In the second section, we discuss the role of  $\underline{h}$  in effecting qualitative vowel changes in Hindi, i.e., the change of  $\underline{a}$  to  $\underline{e}$  and that of short high vowels to their corresponding mid variety. The third section discusses the nature of  $\underline{h}$  from both the traditional phonetic and the generative points of view. We conclude that although the nature of  $\underline{h}$  as given by SPE is able to explain some facts regarding the behaviour of  $\underline{h}$ , the traditional phonetic analysis of  $\underline{h}$  as a fricative also has its strong points. Finally, we propose that  $\underline{h}$  be treated as the non-syllabic counterpart of the low back vowel  $\underline{a}$  rather than of the low front vowel  $\underline{æ}$  as does SPE.

Chapter Three has two main sections : the first on quantitative vowel alternations and the second on vowel nasalization. The first section discusses the various contexts in which vowel length is predictable in native Hindi words either phonetically or morphologically. Phonetically, the vowel is always long word - finally, and always short when followed by two consonants. Morphologically, the root or stem vowel is always shortened when a derivational

suffix is added to a verbal root or a noun/adjective stem. We then try to show that the apparent cases of qualitative vowel alternations can also be subsumed under quantitative ones. In the second section, we take up the predictability of vowel nasalization in Hindi and conclude that although in most cases vowel nasalization is predictable, there are some recalcitrant cases which require an underlying [+ nasal] feature.

In Chapter Four, we discuss the underlying representation of the high glides y, v and the diphthongs ai, au. We have shown that the glides need not be represented as such in the underlying representation, but can be derived from high vowels followed by other vowels; similarly the diphthongs occur as vowel sequences ai, au underlyingly, and the non-syllabic off glide can be again derived by another rule. We take up the question of diphthongs again in Chapter Five.

Chapter Five mainly deals with the underlying representation of the mid vowels. Three possibilities are considered : first, in view of their alternation with the short high vowels in both nominal and verbal paradigms, they can be derived from the latter; second, they can be underlyingly mid vowels; third, they can be derived from underlying vowel sequences, i.e. ai for ē and au for ō. We show that this



last alternative is able to explain a large number of synchronic and diachronic facts about OIA and NIA languages and dialects, which the other two are not capable of doing. As there are no facts which the last analysis cannot explain but the other two can, we accept this analysis. Towards the end of the chapter we show the relationship of the rules deriving glides, diphthongs, and mid vowels, as all these rules act on underlying vowel sequences.

Chapter Six deals with the deletion of the word-final short a. The question of positing a short a word - finally for all those words which phonetically end in a consonant is different from that of positing a short a medially. We claim that despite the absence of alternate forms with word-final a, there is a lot of evidence to justify the positing of an underlying word - final a in Hindi. In particular, a word-final a makes it possible to give a uniform account of all occurrences of flaps deriving them from intervocalic voiced retroflex stops. We also find that if the rule is formulated as a short low back vowel deletion rule, it can become a very general rule applying in Eastern NIA languages also, languages which do not have a short a phonetically. In fact, it turns out to be an NIA rule.

Chapter Seven discusses the question of Hindi syllable structure very briefly. Although, no solutions in terms

of syllable-structure are offered to the problems we have dealt with, our discussion of various rules clearly establishes the relevance of syllable structure to a number of them. In fact, at times, it appears that syllable structure is the motivating factor behind these rules. The brief observations made in this chapter are offered more as indications of the direction in which answers to many unresolved problems of Hindi, indeed NIA, phonology may be found rather than as definitive solutions.

The eighth chapter brings together the main conclusions drawn from our investigations and indicates areas of relevance for future research in the field of Hindi and NIA phonology. The Appendix contains a list of all the rules proposed by us in their final form.

## Chapter 1

### INTRODUCTION

#### 1.1 PRELIMINARIES

If we date the arrival of Transformational Generative Grammar (TGG) on the linguistic scene from 1957, the year of the publication of Syntactic Structures, it has been with us for a little more than twenty years. Much work, both in theory and the description of individual languages, has been done during this period in this framework, in syntax as well as in phonology. The theory has been revised twice by Chomsky himself, first in 1965 to give the Standard Theory (ST) of Aspects, and then in 1971 to give the Extended Standard Theory (EST) of Chomsky 1969. During the same period Chomsky's theory has been attacked by Postal, the Lakoffs, McCawley, and Fillmore, among others. In generative phonology, on the other hand, the paradigm (in Kuhn's sense) was established by the publication of The Sound Pattern of English in 1968 by Chomsky and Halle (SPE). After this, though there have been attempts to revise and modify the framework, notably by Kiparsky 1968a, 1968b, 1971, 1972, Anderson 1969, Coats 1970, Bresnan 1972, Kisseberth 1969, 1970a, 1972, Stampe 1969, the basic framework seems to remain unchallenged and has formed the basis of all work on generative phonology since 1968.

## 1.2 HINDI AND NIA LANGUAGES

In spite of tremendous creative activity in the field of TGG, Modern Indo Aryan (NIA) languages have remained comparatively unexplored, this is more so in generative phonology. In this sense Hindi is no exception. Apart from a few dissertation-length studies like Kachru 1966, Verma 1967, Pray 1970, Sah 1971, Sinha 1972 and Ohala 1972, not many others are available. All of these, except Pray 1970 and Ohala 1972, are on syntax; then Kachru 1966 and Verma 1967 have used the framework of Syntactic Structures. The list of papers also does not fare much better. One of our aims in undertaking this work is to examine, howsoever imperfectly, some aspects of Hindi phonology within the framework of generative phonology.

1.2.2 It is not the case that Hindi phonology has not been studied linguistically. However, as these works have been written from the Neo-Bloomfieldian point of view, they suffer from all those shortcomings to which the framework is subject to; we can specifically mention Dixit 1963 and Kelkar 1968. The latter presents a fairly comprehensive and meticulous account of the phonetics and phonemics of both Hindi and Urdu, and we have referred to this work a number of times.

### 1.3 THE USE OF HISTORICAL AND DIALECTAL INFORMATION: THE NEO-BLOOMFIELDIAN VIEW

1.3.1 One of the basic tenets of Neo-Bloomfieldian linguistics has been the strict separation of the synchronic and the diachronic levels, a separation that can be traced back to Saussure. In the case of Saussure there was enough justification for this separation. It was he who established descriptive or synchronic linguistics in an age when the prevailing intellectual climate favoured the view that the historical study was the only scientific study of language. Saussure's aim was to show that a language could be studied scientifically without making any reference to its history, i.e. by taking the language as it was at any given moment of time. From a practical point of view, there was a similar justification in the case of Bloomfield, Sapir and other pioneers of American structuralism. These linguists were primarily concerned with American Indian languages, for which historical information was totally lacking, and the languages had to be taken as they were. But to elevate this purely practical limitation to the status of a theoretical tenet and to insist that everywhere and in every case this practice should be followed is, in our opinion, totally unjustifiable. We are not claiming that the two, the diachronic and the synchronic, are identical or even non-distinct (in the sense of Chomsky 1965: 181); we are also

not claiming, as Hermann Paul did, that the historical study of a language is the only scientific study. We are simply claiming that the dividing line is not as sharp as it is made out to be and that bringing in historical information can at times help us in understanding phenomena which are purely idiosyncratic in a synchronic grammar; consider, for example, the velar/palatal alternations in Sanskrit. As a language at any given point of time is a product of its history, we think that this fact should be self-evident.

1.3.2 One other important tenet of Neo-Bloomfieldian linguistics is that a given language has a monolithic phonemic system; that is, a language cannot have two phonemic systems. (But see, Fries and Pike, 1949.) It follows from this that every language has its own system of phonemes, all of which are of equal status, and that any division of the system into native and foreign strata, for example, is impermissible. Linguistic studies have amply disproved this contention. (See Chomsky and Halle 1968 for the existence of such a division in English, Postal 1968 for Mohawk, Schane 1968 for French.) And if each language is a self-contained system in this sense, then no use can be made of information from related languages and/or dialects. Thus a grammar of a language is not supposed to give any explanation of the fact that its dialects are mutually intelligible, in spite of differences on the

surface. In other words, it is an idiosyncratic fact about languages. We are unable to accept this position. We are not claiming that there is no difference between the grammar of a language and a comparative grammar of its dialects. We are simply claiming that information from related dialects can help us in understanding the language being examined. We are also not saying that if such information is not available, the linguist should give up writing his grammar and start tearing his hair. Nothing so drastic. But if such information is available and if it can help us, why not use it?

#### 1.4 THE USE OF HISTORICAL, DIALECTAL AND GRAMMATICAL INFORMATION: THE TG VIEW

1.4.1 We thus support the stand taken by generative phonology, or more generally, by TGG, on these issues. Generative phonology does not consider this mutual intelligibility among related dialects as an idiosyncratic fact about languages. Related dialects are mutually intelligible because their differences are normally due to a change in the rule component, rather than to restructuring, i.e., a change in underlying representations. (See King 1969.) Thus generative phonology is able to give a theoretical explanation for something which is accidental and thus unexplainable as well as irrelevant in the Neo-Bloomfieldian framework. It follows from this that evidence from related

dialects can help us in arriving at or deciding about underlying representations and/or the required rules in the language under consideration. In the present work, we have utilised information both from the history of Indo Aryan (IA) as well as from NIA languages and dialects. We have used historical information from IA to show that the cluster approach to aspiration is far superior to the unit segment approach in that it not only explains synchronic facts but also diachronic facts of Hindi. We have made use of facts from the history of IA as well as from NIA languages and dialects to examine the mid vowel system of Hindi. We have tried to show that this way we can arrive at a much more general, though abstract, representation of the mid vowels, which can in turn help us not only in understanding some facts about Hindi and some other NIA dialects but also in establishing a broader unity among these languages. We have tried to do precisely the same thing while discussing the question of a-deletion in Hindi. We have attempted to show that if we take Bengali and other Eastern NIA languages into consideration, we can write the a-deletion rule in a very general form; thus it ceases to be a rule specific to Hindi phonology and becomes part of NIA phonology. We would like to emphasise here that in none of the cases the goal of synchronic explanation or accountability has been lost sight of.



#### 1.4.2 The Use of Grammatical Information

Generative phonology does not consider grammatical information inadmissible evidence for, or irrelevant to, phonological analysis. The exclusion of grammatical information is a legacy of the Neo-Grammarian attitude to sound change. The Neo-Grammarians maintained that the context for sound laws is purely phonetic and that grammatical facts have no effects on sound change. (See Hockett 1965, Postal 1968, King 1969. The last two give examples of sound change where grammatical information is relevant.) This dichotomy between phonology and grammar was extended to descriptive linguistics and enforced most rigorously by the Neo-Bloomfieldians. In this regard, Pike's was a voice in the wilderness. (See Pike 1947, 1952.) On the other hand, the attitude of generative phonology toward this issue is quite different. There are some rules which depend purely on phonetic contexts; there are some other rules which need grammatical information. As such, we cannot totally disregard grammar. We do not say that all rules require grammatical information, but that some rules do.

#### 1.5 THE THEORETICAL FRAMEWORK

The basic theoretical framework of this dissertation is that of The Sound Pattern of English (SPE), which we

have found to be quite adequate for our purposes. We have, however, made some modifications in the distinctive features as given in that work. We have done this as we think that the features that we have used are necessary not only for Hindi but also for other NIA languages, in some cases even for Dravidian languages. These modifications are as follows.

#### 1.5.1 Modifications in the Framework

To describe the retroflex consonants we have used the feature [+ret], i.e., retroflex, and not [+distributed] as the SPE does; we feel that in the case of retroflex consonants, it is the curling back of the tip of the tongue which is important and not the relative area of contact of the tongue with the hard palate. That is, in this case we have accepted the traditional phonetic approach to retroflexion. Even if the area of contact is taken to be relevant, it can be so only in the case of consonants where a contact between the tongue and the hard palate or some other part of the oral cavity is possible. But in some Dravidian languages, we get retroflex vowels as well (Zvelebil 1970: 38). It is clear that in this case the feature [distributed] cannot be used at all, as there is no contact. But the feature [+ret] as signifying the curling back of the tip of the tongue is as relevant here as in the case of consonants.

For the voiced retroflex flap [R] we have used the additional feature [+flap], the feature being equivalent to one quick flap of the tongue against the hard palate. Aspiration in the SPE means the presence of heightened sub-glottal pressure [+hsgp]; we have simply used [+asp]. There is experimental evidence to show that in Hindi the production of aspirated consonants does not involve any rise in the sub-glottal pressure. (See Ohala and Ohala 1972. Halle and Stevens 1971 have also rejected heightened sub-glottal pressure as the cause of aspiration.)

For vowels the SPE uses the tense/lax opposition rather than the long/short one; Pray 1970 has used the same opposition for Hindi also. Although it may generally be the case that long vowels are tense and short ones are lax, we do not think that the matter is one simply of terminological equivalence. In a stress-timed language like English, tenseness can very well correlate with stress. However, in a syllable-timed language, whether Latin, Sanskrit or Hindi, it is the syllable quantity that matters. This syllable quantity is related to the number of moras, or mātrās, in the syllable, which in its turn is related to the length of the vowel; thus long vowels are two moras and short ones are one mora. It is clear that the long/short opposition is much more relevant for Hindi rather than the tense/lax one, as the former can be obviously related to

the number of moras. We have also found that there is a rule in Hindi which shortens the stem or the root vowel of a word when a derivational suffix is added to the word. It seems to us that this is due to a constraint in Hindi governing the number of moras in a word (see Chapter 7). We might add that traditional Indian grammars have used the long/short opposition. We have preferred to follow them in this work rather than follow the tense/lax distinction.

#### 1.5.2 Abstractness versus Concreteness

We have stated that the theoretical framework used in this dissertation is basically that of the SPE with some minor modifications discussed above. However, there is one important respect in which we differ from it. This difference lies in our attitude toward the abstractness of phonological representations. The controversy between abstractness and concreteness of phonology was started in 1968, the year of the publication of the SPE. Kiparsky 1968b set the ball rolling when he criticised Chomsky and Halle for being too abstract in their approach to English phonology. His main objection was against Chomsky and Halle's positing abstract segments like the voiceless velar fricative [x] in the underlying representation of words like right, might, etc., and then deleting them by a rule; that is, his representation was against absolute neutralization, i.e., positing a segment in the underlying representation which

never surfaces. Kiparsky proposed that in the absence of surface alternations, the phonetic shape should be taken as the underlying representation, all other things being equal. He further stated that the Chomskian practice amounts to a diacritic use of a phonological feature; as the theory already provides us with diacritic features to deal with exceptions, this use of a phonological feature for the same purpose is unjustifiable. A number of linguists came out in support of the abstract analysis and produced evidence from various languages in support of absolute neutralization, notable among these being Hyman 1970 for Nupe, Kisseberth 1969 for Yawalmari, Brame 1972 for Maltese Arabic and Vennemann 1970 for Modern German. Although Kiparsky has his supporters (see Kiparsky 1971 for references), the controversy has not been yet resolved. We are inclined to agree with Kiparsky, and consequently, in this dissertation we have not adopted abstract solutions, if a concrete alternative was available, all other things being equal.

A corollary of the above may also be mentioned. Normally what has been done so far in generative phonology is that given a rule X which produces segment A, all instances of A have been derived by X. For example, Chomsky and Halle posit an underlying tense [ī] even for words like kite, high, etc. where there are no alternations of [i] and [ay] as in divine : divinity, and derive the phonetic [ay]

in words of the kite-type also by Vowel Shift Rule. We agree with Vachek 1970 that this derivation is totally unmotivated. Consequently, we have not taken this approach unless it has yielded a more explanatory treatment of relevant data. We may add that the approach taken by Chomsky and Halle, and other linguists as well, seems reminiscent of the once-a-phoneme-always-a-phoneme approach of the Neo-Bloomfieldians in that both accept only one source for a segment.

## 1.6 SUMMARY

The dissertation deals with three main topics. The first topic is that of glides and consists of Chapters 2-3. Chapter 2 centers round the nature of h, which we treat as a low glide, but we also discuss the related question of aspiration at some length. We also take up the question of certain qualitative vowel changes in Hindi and seek an explanation for them in the nature of h. Chapter 3 discusses the high glides, y and ɥ, and the diphthongs ai and au. The second topic is that of vowels and consists of Chapter 4-5. Chapter 4 deals with vowel quantity and vowel nasalization. Chapter 5 deals with the underlying representation of mid vowels. The third topic is that of syllable structure, and comprises Chapters 6-7. Chapter 6 deals with the deletion of word final a and its effect on syllable structure. Chapter 7 has some very brief comments on the

importance of syllable structure in Hindi phonology.

### 1.7 SCOPE OF THE TERM HINDI

Some comments regarding our use of the term Hindi will be in order here. We have taken Hindi as synonymous with Kelkar's Hirdu, i.e., the core common to both literary Hindi and Urdu; another term to denote this common core is Hindustani. The term Standard Hindi has also been used by some writers, e.g., by Dixit 1963. However, we have not preferred this term as, in our opinion, it involves an unnecessary value judgement. We have referred to the various dialects of Hindi by prefixing an adjective denoting the particular region to the word Hindi, e.g., Eastern Hindi (EH), Western Hindi (WH), etc. But we have not taken Avadhi, Braj, etc. as dialects of Hindi, as is the common practice; we have taken them as dialects of NIA, i.e., as sister-languages of Hindi. Although aware of the distinction between Old Indo Aryan (OIA) and Sanskrit, we have used the two terms almost interchangeably, as the distinction is not relevant for our purposes.

### 1.8 CONCLUSION

This dissertation is neither a continuation nor a reflection of earlier work in this field. It is a new work and suffers from all those faults and weaknesses that a new work is liable to. Our aim is not to give a comprehensive

view of Hindi phonology. We have selected certain topics of Hindi phonology which seemed interesting and challenging to us from a theoretical point of view and we have tried to give a generative phonological treatment of these. While doing this we have also tried to show that a generative phonological treatment is superior to an autonomous phonemic one. Since it is more abstract and is not hemmed in by various self-imposed and quite unnecessary restrictions, a generative phonological analysis can capture linguistically significant generalizations to a much greater extent and helps us in understanding and explaining many surface complexities. We do not claim that all such complexities can be so explained; after all, a language does have its idiosyncracies. We simply claim that by taking the help of grammar, historical information about the language, and information from related dialects, we can understand and explain many more facts than we can do otherwise, assuming, of course, that explanation is a concern of the autonomous phonemic approach (see Chomsky 1964a, Postal 1968).

Although our dissertation is on Hindi phonology, we have referred to evidence from other NIA languages and dialects as well. We have done so for the following reasons. First, we believe that evidence from related languages or dialects is relevant in constructing a synchronic grammar of a language. Second, we think that



such evidence can be quite useful in understanding the language under consideration and thus in explanation. Third, our investigations have shown that this explanation is not a one way process; that is, we come to understand both the language and the dialects better. We would like to add that in this way one very interesting fact has been revealed. It is this that a number of rules that we have posited for Hindi are also found in other NIA languages and dialects in basically the same form. For example, we can mention the rule which shortens the stem or the root vowel when a derivational suffix is added to a word and the rule which deletes a, to name only two. Thus, at least these two rules belong to a Pan-NIA grammar rather than to Hindi grammar only. (See Pray 1970 for the idea of a Pan-NIA grammar.) However, as we have not examined any NIA language other than Hindi in any detail, we have confined our remarks and title of the dissertation to Hindi only.

## CHAPTER 2

### ON ASPIRATION AND THE NATURE OF h

#### 2.1 ASPIRATION IN INDO-ARYAN LANGUAGES: PROPOSED TREATMENTS

Indo Aryan languages, whether OIA, MIA or NIA, have been known as languages with a full set of aspirated consonants; as such, aspiration has always been an interesting topic in IA linguistics. The traditional, or structuralist, treatment of aspirated consonants in these languages has been of course in terms of unit phonemes, i.e., the unaspirated and the aspirated consonants have been treated as separate phonemes. Ferguson and Chowdhury 1960 first proposed a cluster analysis for the aspirates in Bengali with a view to reducing the number of phonemes in the inventory, a desirable goal in itself in the structuralist framework. They proposed to analyse aspirated consonants as sequences of the corresponding unaspirated consonant and h. This analysis is equally well possible in the TG framework; we only need a rule of the following type.

TR1: Aspiration

C	h	1	2
		→ [+asp]	∅
1	2		

In fact, both Vajpayee (1977:99) and Kelkar (1968:79) have proposed a similar analysis for Hindi.

However, a particular analysis in generative phonology cannot be taken to be sufficiently justified or well-motivated, if the only aim is a reduction in the number of phonemes. Strictly speaking, the aim has no meaning as generative phonology does not recognise the traditional phonemic level (but see Johns 1969 and Schane 1970). Nor can the justification be that such a rule can be written. The sole justification of a proposed analysis from our point of view is the simplicity of description and the explanatory power of the rule. In other words, we have to see whether the proposed analysis can explain the facts of the language in a simple and principled way or not. The cluster approach, therefore, has also to be examined from this point of view. We feel that a cluster approach can be justified both synchronically as well as diachronically, the criticisms against it by Narang and Becker (1971) notwithstanding. We shall take up these two kinds of justification in the order given; as Narang and Becker's (NB for short) arguments relate to the synchronic aspect only, we shall consider them under that head.

#### 2.1.1 Synchronic Justification for the Cluster Approach

Let us consider those cases where an unaspirated consonant and h come together in Hindi and see what happens.

Take, for example, sequences like rām + hī "Ram only", sēb + hī "apple (emph)", āj + hī "today (emph)", lāl + hī "red (emph)", āp + hī "you (hon) alone", etc. The number of such sequences is very large as the formative hī is an emphatic marker and can thus occur with practically any word. Phonetically, the sequences just given become [rām<sup>h</sup>ī], [sēb<sup>h</sup>ī], [āj<sup>h</sup>ī], [lāl<sup>h</sup>ī], [āp<sup>h</sup>ī] respectively; that is, the unaspirated consonant becomes aspirated and the h disappears as a separate segment. This is precisely what is indicated by our Aspiration Rule given above.

Thus, it is evident that the synchronic facts of Hindi support a cluster analysis. Note that as the aspiration rule that we have given above makes no reference to a boundary, the rule will apply both within as well as across boundaries (SPE: 364). Thus our rule will always give the correct output as long as h happens to be the segment immediately following a consonant. It is to be noted that even in the other analysis where aspirated consonants are posited in underlying forms, we are not able to save ourselves this rule. The rule, or a revised version of it, will have to be in the grammar to take care of the examples given above and many others like them. That is, we shall have two ways of getting the same output. This is not an economical way of doing phonology, if we go by 'economy' as understood in the SPE theory. Thus, we see that the

cluster approach is desirable from both the practical and the theoretical points of view.

#### 2.1.1.1 Criticisms of the cluster analysis

Narang and Becker 1971 have raised two objections against the cluster approach. First, according to them the aspirates behave like unit segments with respect to the a-deletion rule. This rule deletes an a or ə, in certain contexts provided it is not preceded by two consonants. For example, phala 'fruit' becomes phal, but ratna 'jewel' remains ratna, as in the latter case the final a is preceded by an aspirate, the deletion takes place if, of course, the word meets the structural description of the rule otherwise. For example, lābha 'profit' does not remain as it is but becomes lābh. Thus, this behaviour of aspirates is a clear indication of their unit character.

Second, aspirates, like non-aspirates, are represented by a single character in the Devnagari script. Thus, the script treats both of them as units.

Taking the second argument first, the script has single characters for ē, kṣa, tra and jña as well. Do these also become unit segments for this reason? Note that the script is also used in writing Sanskrit, in which ē always comes from ai. Then the aspirated nasals and the aspirated lateral are written as combinations of the

unaspirated consonant and h in Sanskrit, Pali, Prakrits, Hindi, Marathi and all other Indo Aryan languages which use either the Devnagari script or a script derived from it. Finally, the Arabo/Persian script used for writing Urdu represents the aspirates in all cases as a sequence of the unaspirated consonant and h.

As far as the first argument is concerned, although apparently it seems to carry much more weight theoretically, it can also be countered with ease. We have simply to order the a-deletion rule after the aspiration rule, and the objection will be taken care of. Given this proposed ordering, the aspiration rule applying first will convert the C+h sequence into a C<sup>h</sup> unit; after this a-deletion can apply, as there will be no impediment in its way in the form of a cluster. The only way in which the argument can prove conclusive against the cluster approach is to show that the a-deletion rule is the first rule of the grammar, i.e., it is the first rule to apply to the underlying representation. Alternatively, if we can find a situation such that the a-deletion rule has to apply before Aspiration for getting the correct output there, it also could be taken as conclusive proof against the cluster analysis. As neither of these has been shown to be the case, the NB argument cannot be accepted.

Dr. V. S. Khokle (personal communication) has suggested another argument against the cluster approach, which is

interesting as it raises an important theoretical question. According to Khokle the cluster analysis will complicate the underlying phonological structure of Hindi as the sequence CCV is more marked than the sequence CV, because given the first segment as a consonant, the immediately following segment will be unmarked for a vowel but marked for a consonant. (See Ch. 9, SPE.) Now it is a well-attested fact that the development of MIA and NIA from OIA is marked by the process of cluster simplification, which has resulted in a number of CCV sequences of OIA being reduced to C $\bar{V}$  sequences in MIA and NIA. In view of this, how far can the complication of underlying structures in Hindi be justified is a debatable point, especially as an alternative exists which can take care of the relevant facts without this complication. Khokle's point is no doubt well taken. It is true that the cluster approach will make the underlying representations more marked; it is equally true that this approach will ill-accord with the simplification process so well attested in the historical development of MIA and NIA. We turn to a discussion of this problem immediately.

As far as complicating the underlying structure is concerned, what seems pertinent to us is that the theory nowhere says that marked segments and/or sequences are not found in natural languages; it simply says that there is

a cost attached to them. Every natural language has its own quota of costly rules, segments, and sequences. Of course, this does not mean that the linguist should not try to avoid such costly features if possible. But we cannot reject a proposed analysis solely on the basis of its being costly; we have to evaluate its explanatory power as well. We submit that the cluster analysis is much more explanatory than the unit segment one both on synchronic and diachronic counts. It is better synchronically as it explains the aspiration of a consonant when it is immediately followed by h. It is better diachronically as it explains the development of the aspirated consonants from OIA to MIA and NIA (see 2.1.2). It also explains the operation of Grassmann's Law in Sanskrit (2.1.2). None of these is possible through the unit segment analysis. We would like to add that Hyman, who has discussed the question of cluster versus unit segment approaches in general, has left the question open (Hyman 1975: 132ff).

Bringing in the question of historical simplification seems to us to rest on the assumption that surface simplification is equal to a simplification in the underlying representations. While not denying that this may be true in most cases, we do not think that such an identification between the two should be taken for granted. Also,



when it comes to the question of consonant clusters in the underlying representations, at least two types of such clusters will have to be posited there; these are  $C_1C_1$ , i.e., geminate clusters, and NC, i.e., nasal-consonant clusters. In fact, the cluster analysis can provide some motivation for the Aspiration Rule in that the rule simplifies an underlying cluster to a unit segment phonetically. It has to be kept in mind that the rule has got to be there in the grammar in any case to account for the aspiration of a consonant when followed by a formative boundary and h. And the same motivation is there in this case also; thus we are not trying to find motivation for a rule which otherwise is not needed. In fact, both the rule and the motivation for it exist; we are simply extending their scope.

Finally, we would like to add that all the objections discussed above presuppose that h is a consonant, in fact the question of a cluster can arise only if h is a consonant. This, of course, is the traditional phonetic treatment of h. However, generative phonology treats h as a glide, i.e., the non-syllabic form of a vowel, more precisely, of [æ] according to the SPE (but see 2.3 below).

Now glides do not have to occur at the underlying level as glides; they can be derived from the corresponding syllabic vowel by a glide formation rule. (The exact formation of the rule does not concern us here.) Just as phonetic

Cy and Cv (=C<sub>w</sub>) can be derived from underlying Ci and Cu respectively, Ch can also be derived from Cæ. (But see 2.3.3.2 for our criticism of taking [æ] as the syllabic counterpart of the non-syllabic [h].) This, of course, reduces the sequence Ch to an underlying CV sequence, a sequence which is an unmarked one. Strictly speaking, from this point of view, the term cluster analysis is a misnomer as there is no cluster underlyingly. However, as the term is used widely in this sense, we have retained it.

#### 2.1.2.1 Diachronic Justification for the cluster approach

A vast number of aspirates in tadbhavas in Hindi and other NIA languages, as well as in MIA, are not aspirates in Sanskrit tatsamas, but have resulted from the combination of an unaspirated consonant and h/s/ṣ/ś in OIA; the h and the fricatives could be either before or after the consonant in question. For example, Skt. hasta > Hindi hāth 'hand', brahma > brammha 'Creator', aṣṭa > āth 'eight', akṣin > ākḥ 'eye', prastara > patthar 'stone' etc. That is, MIA languages added a rule like our Aspiration rule given above to their grammar. Now we have no intention of saying that a synchronic grammar of a language has to account for facts regarding the history of that language. Yet it is a fact that many historical changes leave their mark in the language in the form of alternations; witness, for example, alternations like line: linear, serene: serenity, etc. in

Modern English, which preserve the Great Vowel Shift. Our claim is that in such a situation the relevant diachronic rule has to be taken as a part of the synchronic grammar of that language. We further claim that this does not mean confusing the diachronic and the synchronic. We shall indeed go a step further and say that a grammar which, without doing violence to the synchronic facts, is able to capture and/or explain historical developments as well is a far better grammar than a purely synchronic one, for the simple reason that it can make us understand language as a developing system in a much better way. We submit that the aspiration of a consonant when followed by h in Hindi is due to the presence of a more restricted form of the same rule which was added to the grammars of the MIA languages. More restricted in the sense that whereas in MIA aspiration could take place irrespective of whether the aspirating segment was h or a fricative, and whether the aspirated consonant was before or after the aspirating one, in Hindi the aspiration takes place only when the unaspirated consonant is followed by h.

Apart from this, evidence can be adduced from Sanskrit to show that the cluster analysis of aspirates can partially explain one fact in Sanskrit which otherwise remains unexplainable; this is Grassmann's Law. According to this law, two aspirates are not allowed to stand in two successive syllables; if two aspirates somehow come to be in such

a position, e.g., through reduplication, the first one is deaspirated. Thus, the root dhā 'to place' becomes in reduplication dadh- in place of \*dhadh-, bhū 'to become' reduplicates as babhū in place of \*bhabhū-, bhī 'to fear' becomes bibhī- instead of the expected \*bhibhī- etc. (These are synchronic cases; we shall mention diachronic examples later on.)

Had the matters stood like this, there would not have been much of a problem; we would have accepted this as a fact about Sanskrit, just as Grimm's Law and Verner's Law are facts about Germanic. After all we cannot give reasons for everything in a language. However, there are certain other facts, again from reduplication, which enable us to look at Grassmann's Law from another point of view. Let us take those roots which have an initial consonant cluster and see what happens in reduplication. For example, we can take śrū 'to hear', pracch 'to ask' sthā 'to stand', jñā 'to know', trā 'to save' etc., which reduplicate like śuśru-, papracch-, tiṣṭh-, jijñā-, tatra- respectively. Even a cursory examination of these forms will tell us that given a consonant cluster, only one of the consonants is chosen for reduplication. A further examination will tell us that the second consonant is chosen if the cluster consists of s/ś followed by a stop, and the first consonant is chosen in all the other cases.

In other words, cluster simplification is the motivating factor behind this behavior. Now let us go back to the case of the aspirates. There it is the non-aspirate which is chosen for reduplication. If we analyse the aspirates as the unaspirated consonant followed by h, their reduplication would be governed by the second case given above. That is, the first consonant will be chosen. And this is precisely what we get. In other words, dhā-: dadhā- is exactly parallel to pracch-: papracch-. Thus we have been able to give a motivation for the application of Grassmann's Law, which otherwise remains without any motivation.

2.1.2.2: The development of PIE voiced aspirates in Sanskrit provides further support for the cluster approach. It is a well-known fact that most of the PIE voiced aspirates have become h in Sanskrit. If we take the cluster approach, then this sound change can be seen as a result of the process of cluster simplification. The reflux of the original voiced aspirate is preserved in certain alternations in Sanskrit. Witness, for example, han 'to kill' and ghnanti 'they kill', duh 'to milk' and dugdha 'milk (n)' etc. In the first case if we take \*ghan as the underlying form, the deletion of g will give us the form of the root, the zero grade of the underlying form will give us ghn-, which is the form of the root in III pl, pr. Similarly, if, in the second case, we posit \*dhugh- as the underlying

form, Grassmann's Law and the deletion of g will give us duh, the actual form of the root; for deriving dugdha, which is really a past participle form, we start with dhugh+ta, Grassmann's Law will give us dugh+ta, voicing and aspiration jump will give us dugh+dha, the reapplication of Grassmann's Law will give us dugdha, the actual form. (As the rules of voicing and aspiration jump are irrelevant to the present discussion, we do not discuss them here.) We are not claiming that this is what actually happens or that if the cluster approach is accepted, all problems in this area of Sanskrit phonology will be immediately solved. We have given these derivations just to show that whereas the deletion of one segment from a 2-segment sequence to get one of the two segments is a comparatively simple process, changing one segment into another by a rule can be a costly affair; when the segments concerned differ as much as gh and h in their feature specifications, the rule becomes very costly indeed. A comparison of the feature matrices of the two segments will make it clear. The feature matrix of g<sup>h</sup> is

$$\begin{bmatrix} - \text{voc} \\ + \text{cons} \\ + \text{hi} \\ - \text{lo} \\ + \text{bk} \\ - \text{cont} \\ + \text{asp} \end{bmatrix}$$

and that of h is,

$$\left( \begin{array}{l} - \text{voc} \\ - \text{cons} \\ - \text{hi} \\ + \text{lo} \\ - \text{bk} \\ + \text{cont} \end{array} \right) ,$$

leaving out the features irrelevant for our purposes. A change of gh to h will necessitate a change of five features, as the values for the features [cons], [hi], [lo], [bk], and [cont] will have to be changed.

2.1.2.3 : In the preceding pages we have considered the arguments given by both NB and Khokle against the cluster approach to aspirates in Hindi and have shown that the arguments given by them in favor of the unit segment approach can in no way be taken to be conclusive proof against the cluster approach. Before proceeding any further we would like to consider an objection to our analysis which is based on phonetic considerations.

We have said earlier that a sequence like āj+hī 'today (emph)' is pronounced as [ā<sup>h</sup>j<sup>h</sup>hī], and thus it rhymes with a word like [mā<sup>h</sup>j<sup>h</sup>hī] 'boatman'. The interesting fact, however, is that whereas the former can be pronounced as āj+hī, i.e., without the aspiration of j, in a slow or careful style the latter is always pronounced as [mā<sup>h</sup>j<sup>h</sup>hī], i.e., with aspirated j, and never as \*māj+hī. This proves that the native speaker is conscious of the difference between these two words with respect to the aspirated j.

If this difference can be related to a difference in underlying representation of the two words, we can explain this phenomenon on a principled basis. The argument is that in the first case there are two segments in the underlying representation and they surface as two segments in certain contexts, but in the second case there is only one segment, the aspirated j, and we get only one segment in all contexts.

Apparently the argument seems quite convincing. However, the facts given in support can be equally well interpreted in another way. Note that **although**, in the cluster approach the aspirates in both the cases will be represented as sequences of j and h, in the first case there will be a formative boundary between the two, whereas in the second case there will be no such boundary. We submit that it is because of this formative boundary that we get the pronunciation [ā<sup>h</sup>j+hī] in slow speech. It is well-attested that in slow or careful speech boundaries become phonetically perceptible as a pause. For example, in slow speech a part and apart are distinguished, though not in normal speech. But apart is never pronounced as a+part.

However, it can still be said that we have been able to make out a case for aspiration in those cases only where an unaspirated consonant is followed by a formative boundary and h. In the absence of any examples of those cases when



the unaspirated consonant is followed by h without an intervening formative boundary our extending the rule to cover these latter cases also is an unmotivated extension of the rule. We shall now show that the language does provide examples for this extension, and that in these cases the consonant is aspirated even though there is no formative boundary between the unaspirated consonant and h.

Consider, for example, Sanskrit words like brahma 'creator', vahni 'fire', jāhnavī 'Ganges', āhlāda 'joy', brāhmaṇa 'brahmin', āhlādita 'joyous', jihvā 'tongue', etc. In normal Hindi these words are pronounced as [bramm<sup>h</sup>a], [bann<sup>h</sup>i], [jānn<sup>h</sup>avī], [āl<sup>h</sup>ād], [brāmm<sup>h</sup>aṇ], [āl<sup>h</sup>ādīt], [jivv<sup>h</sup>ā] respectively. (We disregard the changes irrelevant to the present purpose.) It is clear that first, a metathesis rule applies to these words which changes a medial hC sequence to Ch. Second, another rule of medial consonant gemination also applies which yields CC in place of C, though not necessarily in this order. But the important thing here is that the moment the unaspirated consonant and h come together, the consonant is aspirated and the h is deleted. This is precisely what is the effect of our Aspiration Rule. It is clear that in none of these cases there is a formative boundary present between the consonant and h; in fact, it cannot be as h is to start with to the left of the consonant; it is metathesis because of which

it comes to the right. Thus these examples are clear proof of the fact that h can and does aspirate an immediately preceding consonant irrespective of whether there is a formative boundary between the two or not. We might add that in all these cases once the consonant gets aspirated it stays so; that is, the pronunciation is always [ $\check{j}ivv^h\bar{a}$ ], for example, and never [ $\check{j}ivv+h\bar{a}$ ]. Thus these words are exactly like the word mājhī 'boatman' given above in that in neither case a pause is perceptible between the consonant and h. The only other possibility is that of a formal pronunciation in which no metathesis, and hence no aspiration, takes place; the words are pronounced as they have been written, e.g., [ $\check{j}ihv\bar{a}$ ]. Thus the phonetic facts also are not against the cluster analysis of Hindi aspirates. This case is exactly parallel to the Hindi case given above. In our case also it is the presence or the absence of the formative boundary which decides the issue.

## 2.2 h AND QUALITATIVE VOWEL CHANGE

The role of h in changing the quality of a to e and of u to o is well-attested in Hindi; in fact the former change is a characteristic feature of the western branch of NIA, and the latter belongs typically to Hindi and opposed to Urdu, which maintains u. Although the changes are qualitative vowel changes, they are occasioned solely

by the presence of h, and hence it is not inappropriate to discuss them here. We shall take up the change of a to e first, and that of u to o next.

### 2.2.1 The Change of a to e

The change of a to e in the context of h is a characteristic feature of the western branch of NIA. As Hindi is based on a western dialect of NIA, viz., Khari Boli (=KB), it also shows the change; this fact sharply distinguishes Hindi from Eastern Hindi (=EH), which does not show this change. We might mention in passing that the changes remind us of the behavior of H, or the laryngeal, in PIE. Let us now examine some data.

Consider, for example, words like mahatī 'large (f)'. sahasā 'suddenly' (both from Sanskrit), kahanā 'to say', gahanā 'ornament', mahatō 'a title', rahatē 'living (m,pl)', bahan 'sister', (all native), mahal 'palace', bahas 'argument', sahar 'morning', šahar 'city', (all Arabo-Persian) etc. In EH these words are pronounced as they have been written; for this reason the representations given can be taken as roughly their underlying representations. But in Hindi these words are pronounced as [mahetī], [sahesā], [kahenā], [gahenā], [mahetō], [rahatē], [bahen], [mahel], [bahes], [saher], [šaher] respectively. The h itself may be considerably weakened

or even deleted, and in Western Hindi (=WH) is almost regularly deleted. The deletion of medial h is again a characteristic of the western and northern branches of NIA; we can also compare Modern English and Modern German.

It is clear from these facts that it is the second or the post-h a that is being changed to [e]; it is equally clear that the conditioning factor here is [h]. We propose the following rule to take care of the change:

PR1: a- to e Rule

$$\left[ \begin{array}{l} + \text{ voc} \\ - \text{ cons} \\ + \text{ lo} \\ - \text{ long} \end{array} \right] \rightarrow \left[ \begin{array}{l} - \text{ lo} \\ - \text{ bk} \end{array} \right] / \left[ \begin{array}{l} + \text{ voc} \\ - \text{ cons} \\ + \text{ lo} \\ - \text{ long} \end{array} \right] \text{ h} - \left\{ \begin{array}{l} [+ \text{ cons}] \\ \# \end{array} \right\}$$

As Hindi has only one low vowel, no further specification in terms of the feature [bk] is needed (we are disregarding the underlying [æ]; see 2.3.3.2).

The application of this rule to a sequence like aha will give us [ahe]; this can become [a<sup>h</sup>e] due to the weakening of h, or even [æe] due to the deletion of h. In both these cases, in normal speech, the result is a diphthong with e constituting its off glide, although in slow speech the e might retain its syllabic character. While discussing glides and diphthongs in Chapter 4, we shall show that the underlying sequence ai also becomes phonetic [æe]; in fact the short non-syllabic e is the

phonetic realization of [y] when it is followed by either a consonant or a word boundary. Thus the loss of syllabicity by e when preceded by another vowel is of a much more general nature than would appear here. A full account of this, however, must wait till Chapter 4.

### 2.2.2 The Change of u to o

Let us now consider the change of u to o in the context of h. There are two important differences between this change and the change discussed in the preceding section. First, the u is before h and not after it; second, the u and the o forms are in free variation in Hindi, but in Urdu only the u forms are to be found. That is, the rule is optional and belongs to Hindi phonology only.

Consider, for example, words like suhānā ~ sohānā 'pleasant (m,sg)', kuharā ~ koharā 'fog', muhabbat ~ mohabbat 'love', śuharat ~ śoharat 'fame', suhāg ~ sohāg 'husband', kuhanī ~ kohanī 'elbow (n)', suhabat ~ sohabat 'company', etc. In all such cases it is the presence of h which conditions the u ~ o alternation, although it is possible that the change is governed by the presence of both h and a, long or short. This inference is strengthened by the fact that the two words which we have been able to find with i after h do not show the alternation; these two words are tuhin 'snow' and muhim 'front'. Such words are very difficult to find and mostly belong to the literary

register of Hindi. The alternation is found in some other words also, e.g., muqām ~ moqām 'place (n)', etc. As we are primarily concerned here with the effect of h on vowel quality, and not with qualitative vowel changes in general, we shall not say anything about these alternations. Another important fact about the alternation being discussed here is that it is confined to words of Arabo-Persian origin. Words of non-Arabo-Persian origin like mūha 'mouth', etc. do not show this alternation. The reason for this could be that when such words were borrowed into early NIA from the speech of the native speakers of Arabic and/or Persian, at that time in their speech the short u and the short o were much closer phonetically, at least from the point of view of the NIA speakers. As the o forms of these words are more common in the dialect, and as both Urdu and careful Hindi speakers prefer the u pronunciation, the inference seems to be apparently justified. The following rule can take care of the alternation:

PR2: u- to -o Rule

$$\left[ \begin{array}{l} + \text{ voc} \\ - \text{ cons} \\ + \text{ hi} \\ + \text{ bk} \\ - \text{ long} \end{array} \right] \rightarrow [-\text{hi}] \quad / \quad - \text{ h } \left[ \begin{array}{l} + \text{ voc} \\ - \text{ cons} \\ + \text{ lo} \end{array} \right] \left[ +\text{cons} \right] \quad \left[ +\text{A/P} \right]$$

[A/P = Arabo/Persian].

It should be pointed out that the u~o alternation is not confined to the words given above and others like them. On the contrary, the alternation is quite common in Hindi. But all these common cases depend upon suffixation for word formation. More precisely, when a derivational suffix is added to a stem or root containing long ō, it is shortened to short o. After this, the short o freely varies with the short u in most cases. For example, when the agentive suffix ār is added to sōnā 'gold', we get sunār~sonār 'goldsmith'. However, in these cases also it is the form with the short u which is preferred in both Urdu and Hindi, especially the standard varieties. A notable point is that the alternation here is quite symmetrical in the sense that in similar circumstances long ē is also shortened and this short e freely varies with short i; for example, ēk 'one', but iktārā~ektārā 'single stringed instrument'. This symmetry is not found in the case of alternations due to h (but see 2.2.3). For this reason we have decided to discuss the present u~o alternations separately here. It is quite possible that historically these cases were also cases of alternation after suffixation. But there is no justification for treating it as such in Modern Hindi. For a Hindi speaker words like kuharā~koharā 'fog', śuharat~śoharat 'fame', suhāg~sohāg 'husband', etc. are words in their own right. Deriving them from some

hypothetical stem, i.e., hypothetical for the Hindi speaker, would be totally unmotivated.

### 2.2.3 Possible Extension of u- to -o- Change

We have said in the preceding section that the u o change is not symmetrical, as the corresponding change of i e under similar conditions does not take place. If we could extend the scope of our u- to -o- Rule to include both the high vowels, viz., u and i, it would be quite desirable. For one thing, the rule would become simpler by one feature as no reference to the feature [bk] would have to be made there. Then, the rule would become much more natural and intuitively satisfying in the sense that normally such processes go in pairs, i.e., both the high vowels are affected. Thus, such an extension, if possible, would be desirable from both the formal and the intuitive points of view.

Unfortunately, facts are not so obliging: we do not find parallel instances of iha freely varying with eha. However, a closer look at data brings an interesting point to our attention. Consider, for example, words like chehrā 'face', dehlīj 'threshold', mehṛā 'effeminate', sehrā 'bridegroom's headgear', mehar 'dowry', mēhḍī 'henna', ehsān 'gratitude', mehṛāb 'arch', dehlī 'Delhi', ehsās 'experience (n)', mehtāb 'moon', ehatīyāt 'carefulness',



etc. (Again, all these words are Arabo/Persian.) We are here concerned with these words as they are pronounced by a Hindi speaker with little or no formal knowledge of Urdu. In all these words we find a short e followed by h which in turn is followed optionally by a. (In fact, h is followed by a in all cases; this a is later on deleted by a rule which is independently needed in the grammar. For a formulation of this rule see Pray 1970:93.) On the other hand, when we try to find words in which a short i is followed by h, we find very few of them; in fact, siharnā 'to tremble' was the only common word that we could find. We might add that the words given above in which e is followed by h are quite common.

It is of course possible to take this non-occurrence of i followed by h as an accidental gap in the phonological system. However, in view of the u~o discussed above, it is quite tempting to take this non-occurrence of i followed by h as being due to the application of a parallel i- to -e rule. Parallel because unfortunately it is not possible to combine both the rules into one rule responsible for the alternation of a short high vowel with its corresponding mid variety in the context of h. The u- to -o rule is optional, whereas the i- to -e rule has to be obligatory; these two rules cannot be combined into one. Thus we shall not be able to write a more general rule affecting both

the high vowels. The i- to -e rule also will be confined to Arabo-Persian (=A/P) words. The rule could be formulated as follows.

PR3: i- to -e Rule

$$\left[ \begin{array}{l} + \text{ voc} \\ - \text{ cons} \\ + \text{ hi} \\ - \text{ bk} \\ - \text{ long} \end{array} \right] \rightarrow [-\text{hi}] / \text{---} [\text{h}] \left[ \begin{array}{l} + \text{ voc} \\ - \text{ cons} \\ + \text{ lo} \end{array} \right] [+A/P]$$

Finally, we would like to add that as these two rules are of quite limited application, we have not attempted to formulate them rigorously. Moreover, the second rule, i.e., the i- to -e rule, is quite tentative.

### 2.3 THE NATURE OF h

We have shown above (2.1) that the aspiration of a consonant in Hindi can be the result of an immediately following h; we have also formulated the rule required for aspiration, i.e., TRI. However, in generative phonology, a rule which is natural, in the sense of Postal 1968, Chomsky and Halle 1968, Schane 1972, is more highly valued than one which is not. As a natural rule is one that captures a process which is found in several languages rather than only one, this is not surprising. For example, a rule nasalising a vowel in the context of a nasal is more natural than one which, say, nasalises a vowel before

a fricative; the former is found in several different languages, whereas the latter is not. Formally speaking, we can say that as a nasal is [+nasal], the nasalisation of a vowel means its acquiring this feature; between a fricative and a nasalised vowel there is no such common feature. In other words, rules of assimilation are one kind of natural rules.

Thus if we could find some common feature between h and an aspirated consonant, we could say that it is a natural rule. In our case it would be especially important as the rule, in one form or the other, has always been active in Indo Aryan. But this is not all; that is, we are not concerned with just this rule. We have seen that in Hindi h not only aspirates a consonant but also fronts and raises the low back a to the mid front e as well as lowers u to o, and possibly i to e. Thus our problem would be to find, or at least attempt to find, a common factor in the cases of the vowel change also. In the following part of this section, it is this larger issue of the nature of h that will concern us.

### 2.3.1 Phonetic Theory and h

According to modern phonetic theory (non-SPE), h is a voiceless glottal fricative, although the voiced variety is also possible. According to Panini and Indian grammarians,

h is a voiced glottal fricative; Panini's term for the place of articulation is kanṭhya 'of the throat'; however, as the velars and a have also been described as such, the term most probably refers to backness of articulation or simply [+bk]. The voiced character of h in Sanskrit and NIA languages has been accepted by every phonetician. The SPE, on the other hand, describes h as a low glide; more precisely, it is

$$\left[ \begin{array}{l} - \text{voc} \\ - \text{cons} \\ - \text{hi} \\ + \text{lo} \\ - \text{bk} \\ + \text{hsgp} \end{array} \right] . \quad (\text{SPE:307}). \quad \text{In Hindi}$$

and other NIA languages it will be [+voi] also. The feature [hsgp] stands for [heightened sub-glottal pressure], the SPE equivalent of our [asp]. Thus it is clear that there is no unanimity regarding the character of h; in fact, there is considerable divergence.

### 2.3.2 Aspiration and h

As far as aspiration is concerned, matters are fairly simple if we follow the SPE theory. As aspiration is equivalent to the presence of [+hsgp], the aspiration of a consonant when followed by h, which is also [+hsgp], becomes a case of assimilation. (However, see Halle and Stevens 1971, where this feature has been rejected.) In our case, both h and aspirated consonants will be [+asp], so there will be no problem. We might add here that [+asp]

is equivalent to an extra puff of air during the articulation of the consonant; in other words, we have accepted the traditional phonetic explanation of aspiration.

On the other hand, in the traditional phonetic theory no such common feature is formally available between h and an aspirated consonant; thus it is unable to explain the aspiration of a consonant by h formally. However, we think that traditional phonetics can explain this aspiration in a way which seems intuitively correct, although we are not aware of anyone advancing this explanation. According to traditional phonetics, h is a fricative. Now fricatives are characterised by a strong stream of air passing through the oral cavity, and aspiration involves an extra puff of air. Hence, if a consonant is adjacent to a fricative, this extra puff of air will always be available. Whether it is used or not is a different question. Thus, such a consonant can always be aspirated. This will also explain the aspiration of consonants in Middle Indo Aryan (=MIA) and NIA by  $\underline{s}/\overset{v}{\underline{s}}/\underline{s}$  also, as these are also fricatives. Thus, in both cases aspiration can be explained in terms of a feature, or quality, common to both the aspirated consonants and h as well as the other fricatives. But in generative phonology no such common feature is available between aspirated consonants and the fricatives like  $\underline{s}/\overset{v}{\underline{s}}/\underline{s}$ . This is clearly a drawback from the point of view

of IA linguistics. We shall try to remedy the situation in 2.3.4.

### 2.3.3 Qualitative Vowel Change and h

Let us now consider the change of vowel quality by h, i.e., the change of a to e and of u to o, and of i to e. Formally speaking, the change of a consists in  $\begin{matrix} V \\ [+lo] \\ [+bk] \end{matrix}$  becoming  $\begin{matrix} V \\ [-lo] \\ [-bk] \end{matrix}$ , and in the case of high vowels,  $\begin{matrix} V \\ [+hi] \end{matrix}$  becoming  $\begin{matrix} V \\ [-hi] \end{matrix}$ . Apparently, the change seems to be without any motivation, and from the point of view of traditional phonetic theory, whether Indian or western, it is clearly so. After all why should a glottal fricative, voiced or voiceless, change vowel quality in this manner? We have to bear in mind that this question is relevant not only for Hindi phonology but also for PIE phonology where the role of H, or the laryngeal, in effecting qualitative vowel changes is well-attested. Again, we find that the SPE theory is able to explain things in a formal and principled way, whereas the traditional theory is not. According to SPE (p. 307), h is  $\begin{bmatrix} -hi \\ +lo \\ -bk \end{bmatrix}$ ; we disregard here the features irrelevant for the present discussion. Thus, what happens in our case can be seen as a case of partial assimilation. That is, a, which is  $\begin{bmatrix} -hi \\ +lo \\ +bk \end{bmatrix}$ , becomes  $\begin{bmatrix} -lo \\ -bk \end{bmatrix}$ , an assimilation concerning the part of the tongue raised, and the high vowels become  $[-hi]$ , an assimilation of a feature of tongue height; in both cases the resultant vowel is mid.

This, however, is not the end of the matter; the inadequacy of the traditional phonetic analysis of h, whether Indian or western, is shown in a much more important way. The changes here regarding tongue height and the part of the tongue raised are along two parameters of vowel classification. Consonants, on the other hand, are classified according to the place and the manner of articulation. So if a consonant changes another segment, the change should be along these two latter parameters. Of course, the much more important objection is that normally it is vowels which affect consonants and not the other way round. And even when consonants do affect vowels, e.g., the nasalisation and retroflexion of vowels in the context of nasal and retroflex consonants respectively, qualitative vowel changes are not normally involved. We know of only one language, French, where nasalisation of a vowel results in a qualitative change as the nasalised vowel becomes [+lo]. If h is a fricative, it is a consonant, and so its role in changing vowel quality remains unexplained. But according to the SPE framework, all segments, whether vowels, consonants or glides, have certain features in common, e.g., [bk], [hi], [rnd], etc.; this takes care of the assimilation in our case. More importantly, the claim of the theory that features like [hi] and [bk] are relevant in the case of consonants as well is amply substantiated by

the cases given here. Thus we find that although the aspiration of a consonant by h can be explained by both the SPE and the traditional phonetic theories, the qualitative vowel changes can be explained by the SPE theory alone.

#### 2.3.4 Aspiration and the Fricatives

We have shown (2.3.2) that generative phonology can explain the aspiration of a consonant by h in a principled way as an assimilatory process. This is not possible in the traditional phonetic theory unless it is extended in the way in which we have extended it. However, the aspiration of a consonant by a contiguous fricative in MIA and NIA cannot be explained by generative phonology as it stands. In this section we propose to examine this problem. In our case the question does not involve h, as we have not treated it as a fricative. However, although generative phonologists generally have taken h as a glide, Sloat et al (1978:32) treat it as a glottal fricative; thus in their case the question will be relevant for h also.

The aspiration of a consonant by a contiguous s/š/s is very important in the development of both MIA and NIA. A large number of aspirates of NIA were represented in OIA as sequences containing an unaspirated consonant and a fricative either before or after it. Consider, for example, Skt. aṣṭa > Hindi āṣṭh 'eight', Skt prastara > Hindi patthar



'stone', Skt akṣin > Hindi ākḥ 'eye', Skt paścīma > Avadhi/Braj pačchim 'west', Skt puṣkara > Hindi pōkhar 'pond', Skt pakṣin > Hindi pānchī 'bird', Skt matsya > Bengali māchh 'fish', etc. (Changes irrelevant to the present discussion have been disregarded.) Obviously the problem here is one of finding some common element between h on the one hand, and the fricatives on the other, in terms of distinctive features, so that the aspiration can be attributed to it. As far as the traditional phonetic theory is concerned, such a feature is available. There h is a fricative and all fricatives involve the escape of the air with force or turbulence. This could provide the extra puff of air required for aspiration. But generative phonology as it stands has no such solution, as the feature [+asp] cannot be posited for fricatives. On the other hand, the phonetic explanation given above seems intuitively satisfying to us; consequently, we shall try to incorporate it into our framework.

We think that matters can be improved if instead of looking for one cause of aspiration, i.e., a feature which should be present in both h and the fricatives, we try to see whether aspiration itself can be attributed to more than one cause. We feel that here the traditional phonetic theory, more precisely, the relation between fricatives and aspiration that we have given according to it, can

provide a valuable insight. We have referred earlier to the turbulence in the air stream during the articulation of fricatives. Strictly speaking, this turbulence is not found in all fricatives, but only in a sub-set of them; these are the fricatives which have the feature of stridency, i.e., [+str]. And the three fricatives that we have mentioned, i.e., s, š and ś, are all [+str]. We suggest that [+str] fricatives can also cause aspiration of a contiguous segment. If this is correct, then the aspiration of segments by these three fricatives in MIA and NIA can be explained. That is, in the case of h, the aspiration is due to the feature [asp], whereas in the case of s/š/ś, it is due to the feature [str], all these three being [+str]. But if h is taken as a glottal fricative, it is still left out; however, as we are not affected by this, we shall leave the matter as it stands.

We are quite aware of the tentative nature of the hypothesis put forward; we admit that we can give no experimental evidence in support. It can be said that if stridency of the segment is the deciding feature, then what about the affricates [č] and [ǰ], which are [+str] and found in both MIA and NIA as well as OIA? Why have they not caused aspiration? As an answer to this we can say that it is the stridency of the fricatives that we have been talking about and not stridency in general; that is, it is  $\begin{bmatrix} +cont \\ +str \end{bmatrix}$

that is required and not just [+str]. In fact, the only other fricatives that are [+str] are [f] and [v], and neither of these is found in the languages we are concerned with. More precisely, we can say that the relevant feature matrix for the present purpose is  $\left[ \begin{array}{c} [+cont] \\ \left\{ \begin{array}{c} [+asp] \\ [+str] \end{array} \right\} \end{array} \right]$ , i.e., either h or the strident fricatives.

### 2.3.5 The Underlying Representation of h

The question of the underlying representation of h is closely linked to its treatment as a glide. Glides are non-syllabic counterparts of the corresponding syllabic vowels; thus [y] is non-syllabic /i/, and [w] is non-syllabic /u/. According to SPE, h is a glide having the feature specifications, among others,  $\left[ \begin{array}{c} +lo \\ -bk \end{array} \right]$ ; that is, it is non-syllabic /æ/. It is this analysis that we would like to examine more closely as we think that it has some important theoretical implications that have not been realised fully. In fact, we are not aware of any work using the SPE theory that has discussed the nature of h in any detail.

The theory of markedness as proposed by Chomsky and Halle (1968:Chapter 9; also Postal 1968:Chapter 8) has become a very important part of the theory of generative phonology. The theory incorporates and improves upon the

insights gained by Prague School phonologists regarding phonological structure, especially Trubetzkoy and Jakobson. (Trubetzkoy 1968:27, Jakobson 1968:51ff.) One important aspect of the theory is that the presence of a marked segment in a language presupposes the presence of the corresponding unmarked one, although not vice versa. For example, the presence of voiced obstruents in a language presupposes the existence of voiceless ones, as voicing is a marked quality in obstruents. Taking up the question of vowels and glides, we can say that these two classes of segments can be differentiated solely on the basis of syllabicity, or [syll]; that is, vowels are [+syll] and glides are [-syll]. Further, we can say that a vowel does not have to be positively specified for syllabicity; the marking conventions will automatically make it [+syll]. (Although Chomsky and Halle (1968) do not have any convention for marking syllabicity in vowels, this is what follows from (67) on p. 354.) It follows from this that the presence of non-syllabic vowels, i.e., glides, presupposes the existence of the corresponding syllabic vowels as the latter are unmarked segments as far as syllabicity is concerned. More precisely, the presence of [y], [v] (= [w]), and [h] presupposes the presence of /i/, /u/, and /æ/ respectively at the underlying level.

Now as far as the presence of /i/ and /u/ is concerned, there is no problem; there is hardly any language

in which these two vowels are not found underlyingly, with perhaps the possible exception of Kabardian (Halle 1970). What is more, in many languages, Hindi included, glides can be derived by a rule from the corresponding syllabic vowels, i.e., [y] from /i/ and [v] from /u/; in other words, glides do not have to occur at the underlying level at all. If so, [h] can be derived from /æ/. Obviously, for this /æ/ has to occur at the underlying level. But the important point here is that though [i] and [u] are fairly common or unmarked vowels, [æ], on the other hand, is not so. Although Chomsky and Halle have suggested /a i u æ/ to be the optimal four-vowel system, later on they write that "We are not certain that the system /a i u æ/ occupies the privileged position that we have suggested it does" (1968:410). There are many languages in which /æ/ does not occur; there are several others in which it is derived from some other vowel or vowels. For example, in WH it can be derived from underlying /ai/; in German, in many cases, it can be derived from underlying /a/. But the occurrence of [h] is fairly common in the languages of the world; in fact it is one of the most common segments. This can only lead us to the following two conclusions:

1. In all those languages in which [h] is found, /æ/ is also found. If /æ/ does not occur phonetically, it means that all instances of it have been changed to [h].

CENTRAL LIBRARY

Acc. No. A 63064

In other words, most languages of the world have at least one segment at the underlying level which never surfaces phonetically.

2. The implicational relationship between marked and unmarked segments should be re-examined. We should say that there are two kinds of marked segments. The presence in a language of the segments of the first type presupposes the presence of the corresponding unmarked ones in that language; the presence of the segments of the second type does not. Thus, in the present case, [y] and [v] will belong to the first type, and [h] will belong to the second, as its presence does not imply the presence of /æ/.

If we accept that /æ/ must be present underlyingly in all those languages in which [h] is found, we are in effect claiming that the maximally unmarked vowel system is not /a i u/ but /a i u æ/ for all practical purposes, as it is the latter which would be found in a majority of cases and not the former. We might try to avoid this by positing /h/ underlyingly irrespective of whether we posit -- /y v/ also underlyingly or whether we derive them from high vowels. No doubt, this solution will get rid of the objections regarding the underlying vowel systems. But the fact will remain that the symmetry that we get between

i/y and u/v will be lost as there may be no æ/h; we submit that this will happen in the case of a large number of languages.

If we accept the second conclusion, we are still not out of the woods. The whole idea behind the theory of markedness is to explain in a principled way why in natural languages some segments are more common than certain others, why some segments are not found unless some other segments are also present, why in language acquisition by children some segments appear before others, etc. Once we accept that there can be two kinds of marked segments, we have to concede the possibility of two kinds of unmarked segments as well, two kinds of natural rules, etc. We think that the explanatory power of the theory would suffer this way. In fact, our theory of markedness would become quite like the stratificational theory of markedness where marked/unmarked depend upon a specific language and any segment or component can be taken as unmarked in a given language. For example, Lockwood 1969 takes [ə] as the unmarked vowel in Bulgarian.

Thus we find that treating h as the non-syllabic counterpart of æ leads us into no end of difficulties. To be sure, we can get round these difficulties by making some adjustments in the theory; however, such ad hoc solutions never really touch the core of the problem. The

basic drawback here, as we see it, is that whereas h is one of the most common segments in the languages of the world, æ is not. So if h can be paired with some other vowel, say a, which is equally common, these difficulties will not arise. Needless to say that there has to be some motivation for this. Let us now turn to this question.

The pairing of h with æ is due to the fact that both are treated as being  $\begin{bmatrix} -bk \\ +lo \end{bmatrix}$ , whereas a is  $\begin{bmatrix} +bk \\ +lo \end{bmatrix}$ . Treating h as  $[-bk]$  looks strange from the point of view of traditional phonetics, as h is considered to be a glottal fricative there, and 'glottal' definitely signifies backness of articulation. The justification for treating h as  $[-bk]$  is that the tongue is not retracted from the neutral position during its articulation, but the tongue is not retracted for either  $[\Lambda]$  or  $[\text{ə}]$ ; both these vowels are central vowels in the traditional framework. Then, according to Panini both h and a are kanthya 'of the throat', i.e.  $[+bk]$ . The reason for this difference seems to us to lie in the fact that in generative phonology the feature  $[+bk]$  subsumes both the back and the central vowels of traditional treatments. Thus in generative phonology both  $[\Lambda]$  and  $[a]$  are  $[+bk]$ , but Hockett has treated  $[\Lambda]$ ,  $[\text{ə}]$  and  $[a]$  as both central as well as back vowels (1958:79, Table 9.2).

An obvious corollary of this is that there are bound to be differences in the degree of backness among these



vowels. It is clear that a central vowel and a back vowel are not going to be [+bk] to the same extent; perhaps [-front] will be more explanatory in this regard, especially if we take [-fr] as equivalent to [+bk]. As features are binary, there can be only two values, either + or -. But this applies only to the underlying level; at the phonetic level, features are n-ary, i.e., there are degrees of, for example, voicing, aspiration, backness, etc. Thus at the phonetic level a can be less back or more back, but at the underlying level it will be simply [+bk]. And it is the underlying level that we are concerned with. If we choose to pair h with a, it can be treated as [+bk] with the same distinction regarding the degrees of backness in mind. In other words, h does not have to be phonetically as back as a really back a; it can be paired with a less back, or central, [a].

Although we have tried to show that the pairing of h with a is possible, we have to admit that it can be paired with æ also, just as the SPE does. As both the solutions are possible theoretically, we have to find some other reasons for preferring one solution to the other. One very important advantage of taking h as the non-syllabic counterpart of a is that this way we can explain the high frequency of h in the languages of the world, as a is equally frequent, possibly more so. This is irrespective of whether

we choose to posit h underlyingly in a given language or whether we want to derive it from an underlying a. Then, it can offer an explanation for 'h-intrusion' in some varieties of English. It is to be noted that this intrusion of h takes place only after a, thus cart is pronounced as [kaht]. From the point of view of English phonology, it is perfectly symmetrical in the sense that just as [y] appears after [i], e.g., beat is pronounced as [biyt], [w] appears after [u], e.g. boot is pronounced as [buwt], similarly [h] appears after [a] and [cart] is pronounced as [kaht]. In other words, the intruding or the inserted segment is the non-syllabic counterpart of the preceding syllabic vowel. If h is paired with æ, then this symmetry is lost. Another example can be given from Hindi. In Hindi film songs, which are examples of most popular music in India, a long ā is many times resolved into the sequence ah, but other long vowels do not normally receive this treatment. Again, if a and h are paired together, as we have suggested, this begins to look quite natural. Relating this ā/ah to compensatory vowel lengthening will not do for the simple reason that this is a case of shortening and not of lengthening. In compensatory vowel lengthening the deleted segment is present underlyingly and it can be any consonant, not just h.

## CHAPTER 3

### ON GLIDES AND DIPHTHONGS

#### 3.1 ALTERNATION OF HIGH VOWELS AND GLIDES

The alternation of a high vowel with the corresponding glide, i.e. i ~ y, u ~ v (=w), is quite well known in OIA; the alternations, in fact, go back to PIE. (See Edgerton 1943.) In OIA these alternations are both phonetic and morphological. By phonetic alternations we mean that a high vowel becomes a glide when it is followed by another vowel, e.g. nay-a from nai-a 'to lead', bhav-a from bhū-a 'to be' etc. Recently, Schane has proposed this rule of glide formation as one of his natural rules; that is, a rule of the following type (Schane 1972).

NR1: Glide Rule (GR)

$$\left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{hi} \end{array} \right] \rightarrow [-\text{voc}] \quad / \quad - \quad \left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \end{array} \right]$$

According to Schane, a natural rule does not have to be stated as part of the phonology of a specific natural language; rather, it belongs to the universal grammar. The idea behind the concept of a natural rule is that a rule or a process which is found in a number of languages

has certainly a different status from that of a rule which is found in only one language. Thus our theory should be able to capture this fact. By making such rules as part of universal grammar, we are doing precisely this. So if we can derive glides in Hindi through this rule, it would mean that our grammar would become simpler by one rule, all other things being equal.

### 3.2 ALTERNATIONS IN DEMONSTRATIVE FORMS

If we take the alternations of the demonstrative forms in Hindi, it becomes clear that glide vs high vowel alternation is purely morphological; it is also clear that it is confined to two forms only. The relevant paradigms are as follows.

1. yeh 'this (dir)' ~ is 'this (obl,sg)' ~ in 'this (obl,pl)'
2. voh 'that (dir)' ~ us 'that (obl,sg)' ~ un 'that (obl,pl)'.

It is clear that we get the glides in the direct forms and the high vowel in the oblique ones. The high front glide/vowel goes with the proximate demonstrative, and the high back glide/vowel with the remote one. However, as the alternation is confined to only two items in the language, it cannot have much significance for Hindi phonology in general, whether the alternation is morphological or phonetic. Of course, it is clear that as far as these two forms are concerned, the glides can be derived from the corresponding

direct case for the proximate and remote demonstratives is clearly not general enough. As h is common to both it can be separated. However, what is more important is the pairing of e with i and o with u, i.e., mid front with high front and mid back with high back. Thus, if we can have just a mid vowel having the same value for backness as the preceding high vowel, the process would become much more general. The second point to note is that the pronunciations yeh and voh, alternatively yē and vō, are current in Hindi and Western Hindi. As far as Eastern Hindi (=EH) is concerned, the pronunciations are yah and vah. We think that our solution should be general enough to account for both the pronunciations; this is possible if we represent the direct marker as simply ah, or more precisely, a+h. Then, as far as EH is concerned, i+ah and u+ah will give us yah and vah respectively, as the natural rule of glide formation will be as much a part of EH phonology as of Hindi phonology. The problem now will be to get yeh and voh for Hindi. It is possible to take care of this change by a minor rule for Hindi which would change a to e in the proximate demonstrative and to o in the remote demonstrative. The h can be later on optionally deleted, which then leads to a lengthening of the resultant mid vowel. This minor rule can be stated as follows.

PR4: a-Raising

$$\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{lo} \end{bmatrix} \rightarrow \begin{bmatrix} \alpha\text{bk} \\ -\text{lo} \end{bmatrix} \quad / \quad \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{hi} \\ \alpha\text{bk} \end{bmatrix} + \text{---} + \begin{bmatrix} -\text{voc} \\ -\text{cons} \\ +\text{lo} \end{bmatrix}$$

However, it is possible to do away even with this minor rule. While discussing h, we have shown (2.2) that h changes an immediately preceding [a] to [e], provided no long vowel follows. We have also said there that this change takes place in Hindi and WH but not in EH. Thus, yah will automatically become [yeh], but vah will also become [veh], a form that we do not get. However, for a number of people vē exists as the direct plural of the remote demonstrative. Thus, we would still require a minor rule for getting [voh] from vah. Although the number of rules will not change, we think that this latter solution is a better one as it makes use of an already existing rule and thus makes the situation regarding the pronunciation of glides part of a larger and more general aspect of Hindi pronunciation.

### 3.3 GLIDES IN OTHER FORMS

We would now like to examine whether this analysis can be extended to cover the glides in other words in the language. Although we are hampered by the fact that the kind of alternations that we found in the case of the

demonstratives are not found elsewhere, we think that the pronunciation of y and y in Hindi can provide ample scope for such an extension. We shall be concerned with the high glides, y and y, only and not with the low glide h for two reasons. First, the natural rule proposed by Schane that we have been using specifically mentions a high vowel to the left of the arrow; second, although alternations of high glides with high vowels are quite common, we do not know of any instance where the low glide h alternates with its low vowel counterpart, whatever that is.

We have had occasion to remark earlier (2.1) that in generative phonology the primary justification of a rule is not whether we can make it work, but rather whether it has any motivation in the facts of the language concerned. In this case also it will be the same. With this aim in view, we introduce the following fact about Hindi pronunciation.

In Hindi, the full pronunciation of glides, both y and y, is found when the glides are post-vocalic but not followed by a vowel; elsewhere their pronunciation is like that of the corresponding short non-syllabic mid vowel, i.e. [e] for y and [o] for y. (Phonetically these are extra-short.) This is true of native as well as of Sanskrit words pronounced informally as native words. But the important point is that in slow and careful pronunciation,

the glide always has a full pronunciation. This clearly shows that e and o cannot be underlying but have to be derived from /y/ and /v/. Let us now examine some data. For example, words like yadi 'if', gayā 'went (m,sg)', vahā 'there', nāvik 'boatman', etc. are always pronounced as [yadi], [gayā], [vahā], [nāvik] respectively. In these examples, either the glides are not post-vocalic, or, if they are, they are followed by a vowel; in both cases we always get the full pronunciation. However, words like samaya 'time', gāya 'cow', nāva 'boat', avasara 'chance', etc. are normally pronounced as [samae], [gāe], [nāo], [aosa] respectively, although in slow and careful pronunciation they will be pronounced as they have been written. The diphthongs /ai/ and /au/, both in Sanskrit and Hindi words, are also pronounced as [ae] and [ao] respectively. We are of course talking of the normal, non-formal, Hindi pronunciation. If we represent our glides in the underlying representation as sequences of a high vowel followed by another vowel, we shall be able to offer a very simple explanation for this phenomenon. The GR, which as a natural rule will not belong specifically to Hindi grammar, will turn the high vowel into a glide; after the high vowel has been made [-voc] by GR, we shall require another rule to lower this non-syllabic high vowel to mid. This rule will give us [e o].



We have said earlier (3.3) that the full pronunciation of glides is found only when the glides are post-vocalic but not followed by a vowel, i.e., between a vowel and either a consonant or a word boundary. However, in words like samaya 'time', nāva 'boat', gāya 'cow', avasara 'chance', nayan 'eye', etc. the glide, although post-vocalic, is not followed by either a consonant or a word boundary. Yet the normal pronunciation of these words is [samae], [nāo], [gāe], [aosa], and [naen] respectively, although in a slow and careful pronunciation, they will be pronounced as they have been written, i.e., like [samaya] etc. It is clear from these cases that the post-glide a is being deleted; this deletion results in the glide being followed by a consonant in words like nayan, avasara, and by a word boundary in words like nāva, gāya. (For the deletion of medial a, see Pray 1970: 93; for the deletion of final a, see Chapter 6 of this work.) If the post-glide vowel is a long ā, it is not deleted; thus, in words like gayā 'went (m,sg)', chhāyā 'shadow' etc., there is no deletion. If the post-glide vowel is a short high vowel, either i or u, then again no deletion of the vowel takes place, e.g., in śāyikā 'berth', nāvik 'boatman', bhāvuk 'emotional', āyudh 'weapon', etc., there is no deletion of the high vowel. But the glide itself can be deleted if the glide and the following high vowel have the same value for [bk].

Thus, śāyikā and bhāvuk can become śaikā and thāuk respectively but nāvik and āyudh remain as they are. The rule may be formulated as follows.

PR5: Glide-deletion

$$\left[ \begin{array}{l} -\text{voc} \\ -\text{cons} \\ +\text{hi} \\ \alpha\text{bk} \end{array} \right] \rightarrow \emptyset \quad / \quad - \left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{hi} \\ \alpha\text{bk} \end{array} \right]$$

As far as the pre-glide vowel is concerned, we have seen that if it is a, either long or short, the deletion of the post-glide a takes place, e.g., gāya 'cow', and avasara 'chance' become [gāe] and [aosar] respectively. If the pre-glide vowel is high, no deletion takes place. However, again the glide might be deleted if it has the same value for [bk] as the pre-glide high vowel. Thus, nīyat 'intention', kūvat 'strength', miyād 'time limit', tuvar 'a kind of lentils', etc. are pronounced [nīat], [kūat], [miād], [tuar] respectively, but dhīvar 'a particular caste' is pronounced as it is. The rule responsible for this deletion may be written as follows.

PR5': Glide-deletion

$$\left[ \begin{array}{l} -\text{voc} \\ -\text{cons} \\ +\text{hi} \\ \alpha\text{bk} \end{array} \right] \rightarrow \emptyset \quad / \quad \left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{hi} \\ \alpha\text{bk} \end{array} \right] -$$

We can combine these two rules into one using Bach's Neighborhood Convention (Bach 1968).

PR5": Glide deletion

$$\left[ \begin{array}{c} -\text{voc} \\ -\text{cons} \\ +\text{hi} \\ \alpha\text{bk} \end{array} \right] \rightarrow \emptyset / \left[ \begin{array}{c} +\text{voc} \\ -\text{cons} \\ +\text{hi} \\ \alpha\text{bk} \end{array} \right]$$

Thus it turns out that glide lowering is subject to the glide being immediately followed by either a consonant or a word boundary; this context is provided by the deletion of the post-glide a. We can state the rules as follows.

PR6: Post-Glide a-deletion

$$\left[ \begin{array}{c} +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ -\text{long} \end{array} \right] \rightarrow \emptyset / \left[ \begin{array}{c} +\text{voc} \\ -\text{cons} \\ +\text{lo} \end{array} \right] \left[ \begin{array}{c} -\text{voc} \\ -\text{cons} \\ +\text{hi} \end{array} \right] \text{ —}$$

PR7: Glide Lowering (GL)

$$\left[ \begin{array}{c} -\text{voc} \\ -\text{cons} \end{array} \right] \rightarrow [-\text{hi}] / \left[ \begin{array}{c} +\text{voc} \\ -\text{cons} \end{array} \right] \text{ — } \left\{ \begin{array}{c} [+ \text{cons}] \\ \# \end{array} \right\}$$

The rules will apply in the order given. As h, the low glide, is already  $[-\text{hi}]$ , we do not have to specify  $\left[ \begin{array}{c} -\text{voc} \\ -\text{cons} \end{array} \right]$  further in PR7 in terms of  $[\text{hi}]$ . We give some sample derivations below; we have left out details irrelevant to the present purpose from the underlying representations.

Underlying	iadi 'if'	nāua 'boat'	naian 'eye'
GR	yadi	nāva	nayan
Post-Glide a-deletion	-	nāv	nayn
GL	-	nāo	naen
Surface	yadi	nāo	naen

Thus the set of proposed rules will give the correct output in all cases, no matter what the position of the glide. It is also clear why aya and ava should be pronounced as if instead of two syllables they were only one, i.e., exactly like the diphthongs ai and au. In the diphthongs the second segment, i or u, is non-syllabic; thus they can be represented as [ay] and [av]. As the bisyllabic aya and ava can be reduced to the monosyllabic ay and av, through PR6, the identity of pronunciation is bound to arise. The reason is that as soon as PR6 applies, the SD of PR7 is met and it also applies thereby lowering the glide. This also explains why some Sanskrit words with intervocalic glides have come to have two spellings: one is the Sanskrit spelling, and the other the phonetic one. For example, nayana 'eye' can be written both as nayana as well as naina. However, in formal written Hindi, the phonetic spelling is generally considered incorrect.

## CHAPTER 4

### ON VOWEL QUANTITY AND VOWEL NASALIZATION

#### 4.1 VOWEL QUANTITY

In OIA vowel length was to a great extent predictable either phonetically or morphologically. The phonetic source of long vowels was the sandhi rule according to which two similar vowels, long or short, when next to each other coalesced into the corresponding long vowel. The morphological source was the vridhhi of a under certain morphological conditions, which gave rise to the long ā; for example, from dharma 'religion' we get dhārmika 'religious'. However, if the vridhhi of a vowel, more precisely of a syllabic, is taken as the insertion of an a before the guna form of the vowel, the process can be subsumed under the sandhi rule; thus from dharma we will get dhaarmika, which will become dhārmika by the sandhi rule mentioned above. (The guna of a is a itself.)

When we come to the MIA stage, vowel length becomes still more predictable phonetically, as one more context for predictability has been added. In OIA the vowel length was, generally speaking, not predictable if a consonant

cluster followed the vowel. But in MIA in such a context the vowel is always short; this remains generally true for NIA also. Of course in NIA, especially later NIA, matters are apparently complicated because of large scale borrowing of Sanskrit words as well as of words from Arabo-Persian and European languages. But as far as the native stratum of the vocabulary is concerned, before consonant clusters only short vowels can occur. Thus, in Hindi, vowel length can be predicted in a large number of cases either phonetically or morphologically. We shall discuss these in this order.

#### 4.1.1 Phonetic Predictability

We would like to make one thing clear at the outset of our discussion. We have used the two-fold division between phonetic and morphological predictability here as well as elsewhere purely for descriptive convenience. In generative phonology a rule can refer to grammatical as well as phonetic information, as both are subsumed under phonology. Thus there is no theoretically significant level of morphology. We are using these two terms in a sub-classificatory way only. That is, by phonetic predictability we mean predictability solely in terms of a sequence of segments; by morphological predictability we mean predictability in terms of boundaries, suffixes and other grammatical information.

The sandhi rule of Sanskrit is of course applicable to Sanskrit loans or tatsamas; it may apply to native words also. However, Hindi and other NIA languages have acquired two other rules for predicting vowel length which are absent in Sanskrit, and it is with these rules that we shall be primarily concerned. The first rule, according to which a vowel is shortened before a consonant cluster, has been inherited from MIA. The second rule, according to which a high vowel is lengthened word finally, belongs specifically to Hindi and some other NIA languages. We shall first discuss the rule regarding the vowel length before a consonant cluster.

#### 4.1.1.1 Vowel length before consonant clusters

At times we come across an interesting situation in Hindi, as we get in many cases doublets and/or related words in one of which the vowel is short followed by two consonants and in the other the vowel is long followed by a single consonant. It is clear that in the first case the MIA rule, or constraint, has applied. But what about the second one? We shall come to this question after examining some words exemplifying the two situations.

Consider, for example, addhā/ādhā 'half (m,sg)', matthā/māthā 'forehead (m,sg)', kāndhā/ādhā 'shoulder (m,sg)', āndhā 'blind (m,sg)': ādhi 'duststorm (f,sg)',

hāth(a) 'hand': hatthā 'handle (n)', ḍāṇḍā 'stick (n)':  
ḍār(a) 'punishment', etc. Of course, we are not considering  
 Sanskrit words like śānti 'peace', klānta 'tired', mātsya  
 'of fish', ārdra 'wet', etc. in which a long vowel is fol-  
 lowed by two or more consonants. On the basis of the Hindi  
 words given above, we can say that we get either CVC or  
 CVCC, i.e., strong or heavy syllables, both being preferred  
 syllable structures in Hindi, (see Chapter 7). We propose  
 to deal with such cases in the following way. In the under-  
 lying representation, we represent such words with the  
 sequence of a short vowel followed by two consonants. This  
 will take care of one of the forms. For the other, we have  
 a rule deleting one of the consonants and at the same time  
 lengthening the vowel. This rule can be stated as follows.

TR2: Compensatory Vowel Lengthening (CVL)

$$\begin{array}{ccccccc}
 \left[ \begin{array}{c} +\text{voc} \\ -\text{cons} \end{array} \right] & [+cons] & [+cons] & + & \begin{array}{c} 1 \\ \rightarrow [+long] \end{array} & \begin{array}{c} 2 \\ \emptyset \end{array} & \begin{array}{c} 3 \\ \end{array} & \begin{array}{c} 4 \\ \end{array} \\
 1 & 2 & 3 & 4 & & & & 
 \end{array}$$

There may seem to be some arbitrariness involved in the  
 rule in the sense that it might be said that in the case  
 of geminates we cannot be sure as to which of the two  
 consonants, the first or the second, is deleted. Thus,  
 it is arbitrary on our part to delete the first of the two  
 consonants. The point no doubt is well taken and had our



data consisted of only geminates, there would have been no way to remove this arbitrariness. Fortunately, in our data we also have words with an NC cluster as well as an unaspirate-aspirate cluster, e.g., kāṇdhā and matthā respectively. After the application of CVL to these forms we get kādhā and māthā respectively, i.e., it is the first consonant, the nasal or the unaspirate as the case may be, which has been deleted. So if we delete the first consonant in all cases, our rule becomes much more general. As far as Sanskrit words like śānti 'peace' are concerned, they will be marked in the lexicon as [-CVL] and the rule will not apply to them; thus, we shall never get \*śāti, for example. The reference to the formative boundary is required as the rule applies only within a morpheme and not across it. For example, words like, kattā (=kat+tā) 'is spun (m,sg)', bhartī (=bhar+tī) 'fills (f,sg)', etc. do not become \*kātā and \*bhātī. On the other hand, words like kāttā (=kā+tā) 'spins (m,sg)', dāggghar (=dāk+ghar) 'post office', etc. are not felt to be non-native, although in such words a long vowel is followed by two consonants. Clearly it is the presence or the absence of a formative boundary which is relevant here. We shall see in 4.2.2.2 below that the same constraint is found in the case of nasal deletion.

The problem with this rule, however, will be in another area. It seems that in some cases the rule is

optional whereas in some other cases it is obligatory. We can take care of this situation by taking recourse to lexical marking. But this would mean that this fact is an idiosyncratic property of the word or words concerned. This is actually not so. After examining the data given above, as well as more examples from the language, we can say that the rule is optional in the case of doublets, e.g., matthā/māthā 'forehead'; that is, in those cases in which no derivational suffix is added to the base form. It is a derivational suffix which is relevant here and not just any suffix as in words like matthā/māthā, kāṇdhā/kādhā, etc. the final ā is a suffix marking masculine singular. In words like ḍaṇḍī 'stick (f,sg)', āṇdhī 'blind (f,sg)', etc. the final ī marks feminine singular. In none of these cases CVL applies. On the other hand, in the case of word derivation, i.e., when a derivational suffix is added to the base form, CVL applies, e.g., in ādhī 'duststorm' (f,sg)' < base andh- 'darkness', ḍār(a) 'fine (n)' < base ḍaṇḍ- 'stick', hāth(a) 'hand' < base hatth- 'hand', etc. It is clear that the derivational suffix can have the same phonetic shape as the suffix marking feminine singular, i.e., ī. Thus, the final ī in both āṇdhī and ādhī is the same phonetically, but in the former it marks just the feminine singular whereas in the latter it is a derivational suffix as well. That is, to the base form andh- can be

added the gender number markers ā or ī giving us āndhā/āndhī 'blind (m,sg/f,sg)'. As this is not a case of word derivation, CVL will not apply. But to the same stem we can also add the derivational suffix ī; in this case CVL will apply and give us ādhi 'duststorm'. It seems that the question of the optional application of CVL has arisen because of borrowing. That is, forms like māthā 'forehead', kādhā 'shoulder', etc. have been borrowed from other NIA dialects, e.g., Avadhi. It would appear that in Avadhi CVL applies much more freely than in Hindi; for example, the word for 'blind' is ādhar in Avadhi whereas Hindi has āndhā. Thus we can say that in Hindi CVL is an obligatory rule applying in word derivation, although it is a rule of limited application. Finally, we might mention one other problem in this connection. Although the vowel length has been taken care of by CVL, the nasalization of certain vowels has been left unaccounted for. We find that this nasalization appears only when the vowel is pre-nasal; as such the nasalization becomes explainable on the basis of a very general rule in natural languages according to which a vowel is nasalised when it is pre-nasal.

This universal principle of vowel nasalization may also be used as evidence for the vowel lengthening rule. It could be argued that instead of a vowel lengthening rule we should postulate a vowel shortening rule. In this case

the underlying form would contain the long vowel + a single consonant, i.e.,  $\bar{V}C$ . The vowel shortening rule would add a consonant (double the consonant in the case of geminates) while at the same time shortening the vowel. Thus given an underlying pāt- 'leaf', the rule would give us patt-ā or patt-ī depending on whether it is masculine singular or feminine singular. But this procedure would be unsatisfactory for more than one reason. First, it will necessitate an arbitrary decision on our part regarding the example given and others like it. For example, where do we insert the second t, to the right of the first one or to the left of it? Then, given an underlying māth- 'forehead', the rule would give us \*mathth-ā, a non-existent form with two aspirates in sequence. Using a rule like Grassmann's law would no doubt do the trick, but it would be entirely unmotivated in Hindi. The third problem would be with nasals. That is, given an underlying kādh- 'shoulder', the rule would give us \*kadhdh-ā, a non-existent form again. We might be able to get round this difficulty by putting a condition on the rule that if the long vowel is nasalised, then a nasal will be inserted before the consonant. This again is very ad hoc and devised just to save the rule. Fourthly, the natural connection between the nasalisation of a vowel and the presence of a nasal segment will be lost. In the present account, on the other hand, the

nasal consonant is present in the underlying form and naturally nasalises the preceding vowel. When CVL operates, the nasal is deleted but the nasalization on the vowel stays. This indicates that a vowel lengthening rule presents a more natural account of the facts in the present case than would a vowel shortening rule.

#### 4.1.1.2 Vowel length in word final position

Another context in which vowel length can be predicted is the word final position. In Hindi in this position we get either a long vowel or no vowel at all. We are again not considering Sanskrit words in a formal context. In reality, the cases in which we get no vowel at all finally, i.e., the consonant final words, are all subject to a rule which deletes a word final a. In underlying representation all these words end in a short a. However, as we shall discuss this at some length in Chapter 6, we shall here assume that these words end in a consonant. This final vowel lengthening takes place in the case of those Sanskrit words which end in a short high vowel, i.e., i or u, such short high vowel ending words being absent from native Hindi vocabulary. Consider, for example, Sanskrit words like kavi 'poet', guru 'teacher', sādhū 'saint', sāndhi 'sandhi', śānti 'peace', agni 'fire', kānti 'beauty', viṣṇu 'Vishnu', ravi 'sun', etc., which are normally pronounced as [kavī],

[gurū], [sād<sup>h</sup>ū], [sā<sup>h</sup>ndī], [śā<sup>h</sup>ntī], [agnī], [kā<sup>h</sup>ntī], [visnū], [ravī] respectively. Clearly a rule of the following type is indicated.

PR8: Final Vowel Lengthening (FVL)

$$\left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{hi} \end{array} \right] \rightarrow [+long] \quad / \quad \text{---} \# ] [+native]$$

This rule will lengthen the high vowels only if they occur in word final position in native words or in words being treated as native. This latter qualification is necessary as all the Sanskrit words given above, and others like them, do not undergo this rule if they are being treated as Sanskrit words, i.e., in a formal situation. On the other hand, a number of words given above can be proper names as well, e.g., ravi, śā<sup>h</sup>nti, kā<sup>h</sup>nti, viṣṇu; as such they are always pronounced with a final long vowel. This fact clearly proves that the proposed rule is a native rule. The rule (FVL) will also not apply in the case of those Sanskrit compounds in which such words occur as the non-final member, because the word boundary between such words in a compound is deleted and substituted with a formative boundary by a general rule belonging to the theory rather than the phonology of any particular language. (There may be some exceptions, but these can be treated in the same way as exceptions are generally treated in the theory.)

As examples of such compounds we can take words like kavi-sammēlan 'a gathering of poets', sādhū-svābhav 'saintly nature', śānti-vārtā 'peace talks', etc.

One objection that can be raised here is regarding the length of the high vowel in the underlying representation. It can be said that if we represent the high vowels as underlyingly long, we can save ourselves one rule, FVL; that is, we represent words like kavi 'poet', sādhū 'saint', etc. with long final vowels, e.g., kavī, sādhū etc. However, this will not save us a rule; what it will really result in is that instead of FVL we shall require another rule for shortening the final high vowel in certain contexts, i.e., when it becomes non-final in a compound and when the word is used in a formal context. Thus, as far as the number of rules is concerned, our grammar will not be simpler. Let us now show why it will be more arbitrary and less intuitive.

We have said that the lengthening of the final high vowel takes place in colloquial speech. In a formal context the words given above are pronounced with a short final high vowel. It seems to us that this is a clear indication of the fact that the native speaker is conscious of the fact that such Sanskrit words have a short final vowel and that the lengthening is one way of nativizing the words.

Those Sanskrit words which are rare and not commonly used in Hindi are pronounced with short high vowels if they end in them, showing that in their case there is no nativization. For example, words like vāri 'water', atithi 'guest', ari 'enemy', paśu 'animal', hanu 'jaw', aśru 'tear (n)', ali 'black bee', etc. are pronounced as they have been written, without any lengthening of the final vowel. If to these two facts, we add the third fact that native Hindi words never end in short high vowels, we can safely conclude that our analysis is the correct one and that FVL is needed in a synchronic grammar of Hindi to account for the final high vowel in Sanskrit words in an informal context.

We might mention another fact in defence of our representing words of KAVI type with final short vowels. There are many Hindi words which end in a final long high vowel, e.g., ghōRī 'a mare', bhālū 'a bear', laRkī 'a girl', khiRkī 'a window', ullū 'an owl', ālū 'potato', mālī 'a gardener', etc., let us call these Set 1 words. There are also a number of Sanskrit words which end in final long high vowels, e.g., nadī 'river', vadhū 'wife', dēvī 'goddess', purī 'city', vanī 'speech', etc., let us call these Set 2 words. All these words, of Set 1 as well as Set 2, are always pronounced with final long vowels, and as such they will be represented with underlying long vowels. Now suppose that KAVI type of words also are represented underlyingly with final long vowels and we have a rule for



shortening the high vowels in the appropriate contexts. The problem here will be in explaining the non-shortening of the high vowels in the case of both Set 1 and Set 2 words in precisely those contexts. The only way out will be of lexical marking. But lexical marking signifies an idiosyncratic fact about the form concerned, and no native speaker of Hindi is going to accept the contention that the non-shortening of the final vowels in Set 1 and Set 2 words is an idiosyncratic fact about them. On the other hand, if we accept FVL as part of the grammar, no marking is needed. Obviously, even on purely formal grounds, this is a simpler solution.

We have said that in Hindi words end in either a long vowel or a consonant; the vowels that can occur finally are ī, ū, ā, ē, ō and the diphthongs ai, au, although the frequency of the last four is very small. We are not taking into account inflected words as ē and ō occur as plural markers; that is, we are taking into account only uninflected words. It might be said that if we remove the specification for the feature [hi] from our FVL, we can make it much more general as well as formally simpler. The rule could then apply to all the vowels that can occur in the final position and give us the correct output in each case. Apparently, this seems to be a better solution, but in reality it is not so. First, we have

mentioned above that all those words which end in a consonant phonetically, have a short a finally in the underlying representation. (We discuss this question in some detail in Chapter 6.) This a is deleted by a rule later on. Thus it is clear that there are going to be two types of a-ending words: one in which the a is lengthened, and the second in which the a is deleted. Clearly it is not a happy state of affairs. Second, as far as final high vowels are concerned, we have seen that the KAVI type words shorten their final vowels in certain contexts, i.e., when they are non-final members in a compound and in a formal context. On the other hand, words belonging to Set 1 and Set 2 never shorten their final vowels. This clearly indicates an underlying difference; thus the KAVI type words end in a short vowel and the words of Set 1 and Set 2 end in a long vowel. Now the final long non-high vowels, ā, ē, ō and the diphthongs ai, au, behave just as the long high vowels of Set 1 and Set 2 words do, that is, they also never shorten in the contexts relevant to the present discussion. Thus by the same token, they should also be represented with underlying long final vowels. As far as the KAVI type words are concerned, we have given sufficient justification for final short vowels. But in the rest of the cases there is absolutely no motivation for positing an underlying final short vowel and then lengthening it in every case. Thus

we shall retain FVL as it has been formulated.

The two contexts given above, i.e., the pre-CC context and the word final context, are the only two contexts in which vowel length can be predicted purely on the basis of phonetic configuration. We now pass on to the morphological context for the prediction of vowel length.

#### 4.1.2 Morphological Predictability

The entire context of morphological predictability can be subsumed under suffixal morphology, both derivational and inflectional. Of these two, derivational morphology is much more important as it covers the derivation of both nouns and verbs and as such applies to a large number of words. Inflectional morphology is relevant in the case of words with final long high vowels only. There is a third area of compounding as well, but in Modern Hindi this type of compounding is hardly found, and no new compounds are made on the pattern of existing ones, at least not in formal written Hindi. We shall have something to say about these later on.

##### 4.1.2.1 Derivation of nouns

Consider, for example, alternations like tāl/talaiā 'pond/small pond', khētī/khetihar 'farming/farmer', phūl/phulvārī 'flower/garden', ām/āmraī 'mango/mango orchard', bāṭ/baṭōhī 'road/traveller', kām/kamāū 'work/one who earns',

būjh/bujhakka 'figure out/one who figures out', khīch/khichaiā 'pull/one who pulls', mōṭā/moṭāpā 'fat/obesity', chhōṭā/chhuṭpan 'small/childhood', etc. Before proceeding any further, we would like to mention that the free variation of a short high vowel with its corresponding mid variety is quite frequent in Hindi; this variation generally takes place when a long mid vowel is shortened in word derivation. Thus from mōṭā 'fat', we get both moṭāpā and muṭāpā 'obesity'. We shall have something to say about it in Chapter 5.

It is clear that in the cases given above the shortening of the stem or the root vowel in the second member of the pair is morphological, i.e., it takes place when a derivational suffix is added to the root or the stem. However, it may be said that as the addition of the suffix means the addition of one or two extra syllables, it is possible to subsume this shortening also under phonetic predictability. In other words, the shortening of the vowel is due to some syllable structure constraint governing the number of moras in Hindi words. (See Chapter 7 for some discussion of this point.) Although there seems to be some truth in this point of view, as a justification for our approach we can say that even though the shortening may be attributed to the addition of extra syllables, nevertheless the process is morphological, as the addition of suffixes is a morphological process and

not a phonetic one. Also, although suffixes can be divided into inflectional and derivational, syllables cannot be obviously so divided. This is a very important point as the shortening takes place only when derivational suffixes are added and not when inflectional ones are. For example, there is no shortening of the stem or the root vowel in āmo 'mangoes (obl pl)', phūlo 'flowers (obl pl)', būjhēgā '(he) will figure out', khīchā '(he) pulled', etc., as in all these cases an inflectional suffix has been added. Hence we shall continue to call it a case of morphological predictability. We shall say that the shortening of the long stem vowel takes place when certain suffixes like the diminutive--aiā, the agentive--akkaR, -aiā, -har, -āū, or nominalizers like -ohī, -vārī, -rāī, etc. are added to the stem. (We shall be using the word stem to subsume both stem and root.) We would like to mention again that in a number of cases where we have a long mid vowel in the stem, after shortening, the resultant short mid vowel is in free variation with the corresponding short high one. We might formulate the rule as follows.

PR9: Stem vowel shortening (SV:)

$$\left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \end{array} \right] \longrightarrow [-\text{long}] \quad / \quad \begin{array}{l} \text{X} - \text{Y} + \text{Z} \\ \quad \quad \quad \left[ \begin{array}{l} +\text{der} \\ \text{suff} \end{array} \right] \end{array}$$

#### 4.1.2.2 The derivation of verbs

A Hindi verb can have four morphologically related forms, though not all the verbs exhibit all the four forms. A majority has three, some have only two, and a handful of them have only one. These forms have been semantically defined and named as Passive/Intransitive, Active/Transitive, I Causative, and II Causative. In a number of cases the two causatives have semantically fallen together and there is hardly any meaning difference between the two as far as the actual use is concerned. For example, pilānā and pilvānā, the I and II causatives respectively from pīnā 'to drink' are used interchangeably. Of course, structurally, the difference is maintained and all the four possible forms have different phonetic shapes.

As the structural, i.e., phonetic, difference is still there, it would be better if we could define these four forms on a purely formal basis. This is not to run down semantics or to deny the importance of semantics in linguistics; our consideration in attempting such a classification is a purely formal one. More precisely, we think that such a classification would make the process of verb derivation a part, or rather an extension, of a much more general phonological process in the language. As the semantic classification is correct in most of the cases, we shall refer to it and attempt a correspondence between it

and the structural classification that we shall be proposing.

From a structural, i.e., phonetic, point of view, Hindi verbs can be divided into the following four types.

- I. Weak form, e.g., piṭnā 'to be beaten', kaṭnā 'to be cut', tūṭnā 'to break (intr)' etc. This type consists of either passive or intransitive verbs.
- II. Normal form, e.g., mānnā 'to accept', pīṭnā 'to beat', kāṭnā 'to cut', tōṛnā 'to break(tr)', etc. This type consists of mostly active or transitive verbs.
- III. The -ā-form, e.g., kaṭānā 'to get cut', piḷānā 'to serve a drink', tuṛānā 'to get broken', paṛhānā 'to teach', saṃjhānā 'to explain', etc. This type consists of what are usually called the I causatives. Although some verbs here look like normal agentive or active verbs, yet the causative origin is there. For example, saṃjhānā 'to explain' is the causative of saṃajhnā 'to understand', paṛhānā 'to teach' is the causative of paṛhnā 'to study' etc.
- IV. The -vā-form, e.g., piṭvānā 'to cause to be beaten', kaṭvānā 'to cause to be cut', piḷvānā 'to cause to serve a drink', etc. This type consists of what are generally called the II Causatives. It is between the members of these two last groups that in some cases the semantic differentiation has been blurred.

It is clear that before we are able to account for the various alternations exhibited here, we have to do two things. First, we have to select one of these forms as the base or the underlying form; second, we have to formulate rules which will enable us to derive the rest of the existing forms for any given verb from the base form. As far as the first is concerned, we propose that the forms that we have called the normal forms should be taken as the underlying forms. For this choice we can give the following two justifications. First, all of these forms are active and as such can be taken as more 'basic' than either the passives or the causatives. An apparent exception could be a verb like jalnā 'to burn (intr)'. We propose to take this verb as basic and derive jalānā 'to burn (tr)' as derived from it. We would like to submit that intransitives are not the same as passives and thus can be as basic as the active verbs. Second, by taking these forms as basic, we can simplify the derivation to a large extent.

Coming to the second requirement of the formulation of rules to account for the other existing forms, we submit that no new rules are required. We have already said that one of the plus points of our formalistic classification is that it makes the different alternations exhibited in the verbal paradigm as a part of a very general phonological process in Hindi. This process is the shortening of



the stem vowel when a derivational suffix is added to the stem, a process that we have already encountered in the case of nominal derivation; in fact, the relevant rule (SVS) has already been formulated. We claim that verbal derivation, like nominal derivation, is essentially a suffixal process and that the vowel shortening found in some of the forms in the verbal paradigm can be taken care of by the SVS. We now proceed to show how it can be done.

Consider the alternations exhibited in Table 1. We have listed the various forms group-wise in a linear order, i.e., from I to IV. The gloss gives only the basic meaning. A blank means that no related form exists for that group or if it does, it is not commonly used. The infinitive marker -nā, the marker for the III group -ā- and that for the IV group -vā- have been separated from the stem with a hyphen.

Table 1

I	II	III	IV	Gloss
kaṭ-nā	kāṭ-nā	kaṭ-ā-nā	kaṭ-vā-nā	to cut
-	jal-nā	jal-ā-nā	jal-vā-nā	to burn
-	pī-nā	pil-ā-nā	pil-vā-nā	to drink
piṭ-nā	pīṭ-nā	-	piṭ-vā-nā	to beat
phaṭ-nā	phāR-nā	-	phak-vā-nā	to tear
-	uṭh-nā	uṭh-ā-nā	uṭh-vā-nā	to rise

Table 1 (...Contd.)

I	II	III	IV	Gloss
-	jhuk-nā	jhuk-ā-nā	jhuk-vā-nā	to bend
~sich-nā	~sich-nā	~sich-ā-nā	~sich-vā-nā	to irrigate
-	bujh-nā	bujh-ā-nā	bujh-vā-nā	to quench
-	ḍhūḍh-nā	ḍhūḍh-ā-nā	ḍhūḍh-vā-nā	to search
-	kar-nā	-	kar-vā-nā	to do
-	chamak-nā	chamk-ā-nā	chamak-vā-nā	'to shine'

On the basis of an analysis of these and other such related verb forms in the language, we can mention the following structural points regarding the different members of the series. The consonantal alternations found in some of the forms here as well as in certain other verbs in the language are irrelevant for our purpose here and thus can be disregarded. Class I forms have invariably a short stem vowel; Class II forms may have either a long or a short stem vowel; Class III forms have a short stem vowel as well as a stem-final-ā-; Class IV forms have a short stem vowel and a stem-final-vā. (We might analyse this -vā as y+ā; i.e., we can take -y-as the characteristic suffix for class IV verbs, and -ā as the suffix common to both Class III and Class IV verbs. However, as this detail does not have any direct bearing on the analysis being proposed here, we shall disregard it and continue to treat-vā as a single

suffix.) It is clear that while both Class III and Class IV verbs have an overt suffix, -ā for Class III and -vā for class IV, Class I and Class II verbs have not. As the Class II forms are underlying and as such not derived from any other verb form, the question of a suffix, especially a derivational one, does not arise. On the other hand, the absence of an overt suffix in the case of Class I verbs is a problem. However, as these forms are generally passive, there must be some feature or marker in the deep structure to account for the passive meaning. We feel that this information can be formally expressed by positing a phonologically null passive suffix. Now our derivations will become simple and straightforward. Thus, starting from kāṭ-nā 'to cut (tr)', we can derive the other three forms in the following way.

1. kāṭ+ ∅  $\xrightarrow{\text{SVS}}$  kaṭ(nā) 'to be cut'
2. kāṭ+ā  $\xrightarrow{\text{SVS}}$  kaṭā(nā) 'to get cut'
3. kāṭ+vā  $\xrightarrow{\text{SVS}}$  kaṭvā(nā) 'to get cut'

An absence of a particular form, e.g. Class I or Class III, would simply mean that the verb in question is an exception to that particular derivational suffix, a situation which is amply illustrated in natural languages.

Thus we can account for a large majority of Hindi verbs through one rule and three derivational suffixes. However, before we proceed any further, we would like to

consider some verbs which are somewhat different from the ones that we have just considered. We list these verbs in Table 2 below following the same order of presentation as we did in Table 1.

Table 2

I	II	III	IV	Gloss
nikal-nā	nikāl-nā	-	nikal-vā-nā	to emerge
sāvar-nā	sāvār-nā	-	sāvar-vā-nā	to decorate
sābhal-nā	sābhāl-nā	-	sābhal-vā-nā	to take care
utar-nā	utār-nā	-	utar-vā-nā	to take down
bahar-nā	bahār-nā	-	bahar-vā-nā	to sweep
-	pukār-nā	-	pukar-vā-nā	to call
sudhar-nā	sudhār-nā	-	sudhar-vā-nā	to rectify
uchhal-nā	uchhāl-nā	-	uchhal-vā-nā	to throw
-	baṭōr-nā	-	baṭor-vā-nā	to gather
-	kurēd-nā	-	kured-vā-nā	to scratch

An examination of these verbs brings out four differences as compared to those in Table 1. First, none of these verbs exhibits the full set of four forms. Second, none of these verbs shows the -ā- form, i.e., Class III. Third, in the base forms, i.e., Class II, it is the second vowel which is invariably long whereas in the verbs given in Table 1 it was the first vowel which was long, though not always.

Compare, for example, phāR-nā (Table 1) 'to tear' with nikāl-nā (Table 2) 'to take out'. Fourth, all the verbs given in Table 2 are bisyllabic as opposed to only one in Table 1. Although there are a number of verbs belonging to Table 1 which are bisyllabic, we have given only one of them. But the important point is that all the verbs belonging to Table 2 are bisyllabic. Needless to say the tables are not exhaustive.

Apparently it might appear that SVS will not work in these cases. That is, it would be difficult for the rule to shorten the first vowel in some cases and the second one in certain other cases. We reproduce the rule below for case of reference.

$$\text{SVS: } \left[ \begin{array}{c} +\text{voc} \\ -\text{cons} \end{array} \right] \rightarrow [-\text{long}] \quad / \quad \begin{array}{c} X - Y + Z \\ \quad \quad \quad \left[ \begin{array}{c} +\text{der} \\ \text{suff} \end{array} \right] \end{array}$$

One way out of this difficulty would be to treat the verbs given in Table 2 as consisting of a prefix + a stem, i.e., we analyse nikāl-nā 'to take out' as ni+kāl-nā. Then we constrain the SVS to apply within formative boundaries only. This will take the prefix out of the domain of the rule and only the second syllable, i.e., the first syllable of the stem, will remain for the rule to apply to. No doubt this will give us the correct output, but there are two objections against this approach. First, it will mean complicating the rule. Second, although historically these forms

might be analysed as consisting of a prefix and a stem, as far as Modern Hindi is concerned these forms are single formatives. There is absolutely no motivation for positing a formative boundary within these verbs. Thus this approach cannot be accepted.

However, we feel that there is no need to treat the verbs in Table 2 differently; SVS, as formulated above, is perfectly capable of doing the job. Note that the environment of the rule has an X to the left of vowel to be shortened. Now this X, being the left-hand environment, "may be null, or may consist of units or strings of units of various sorts", (SPE:332). Thus it can be a syllable as well as a segment. Now there is no problem; the rule will take the first long vowel and shorten it whether it is in the first syllable or the second. The only problem that can arise is a succession of two long vowels in a verb, i.e., a long vowel in both the first and the second syllables. As there are no such verbs in Hindi, there is no problem, and the SVS is adequate as it has been formulated above.

Having taken care of the verb forms given in Table 2, we now proceed to another set of verbs, exhibiting qualitative vowel alternations, more precisely, the alternation of long mid vowels with their short high counterparts. Apparently these alternations cannot be taken care of by the SVS as this rule affects vowel quantity only. In fact, the verbs of this

set will have to be marked as exceptions to SVS; otherwise in their case also we shall get quantitative alternations. We list these verbs in Table 3 below, following the same order of presentation as before. The consonantal alternations found in some of these verb forms are irrelevant to the present discussion and will be disregarded. Although the list is not exhaustive, there are not many other verbs like these in the language.

Table 3

I	II	III	IV	Gloss
<u>Group 1</u>				
bik-nā	bēch-nā	-	bik-vā-nā	to sell
ruk-nā	rōk-nā	ruk-ā-nā	ruk-vā-nā	to stop
juḡ-nā	jōR-nā	-	juR-vā-nā	to join
<u>Group 2</u>				
phūṭ-nā	phōR-nā	phuṭ-ā-nā	phuR-vā-nā	to burst
sūkh-nā	sōkh-nā	sukh-ā-nā	sukh-vā-nā	to dry
ṭūṭ-nā	tōR-nā	tuR-ā-nā	tuR-vā-nā	to break

In these cases we find an alternation between a long mid and the corresponding short high vowel, except in the case of the forms listed under Group 2 where in the form under I we get a long high vowel; note also that only the high back vowel u, and of course the mid back ō, is found in the forms in Group 2. Thus these verbs, belonging to

Group 1 as well as Group 2, seem to constitute a separate group of verbs exhibiting qualitative vowel changes as opposed to verbs in both Table 1 and Table 2, which show quantitative vowel changes. However, except in this respect, verbs given in Table 3 do not differ from those given earlier. Thus, forms under I show a short vowel, except in the three cases in Group 2 of Table 3; forms under III always have the short vowel as well as the expected stem final -ā; forms under IV, again, always have the short vowel and the expected stem final -vā. In view of these similarities it would be very desirable if we could find a way to integrate the verbs given in Table 3 with those which we have given in Tables 1 and 2. This would clearly enable us to present a very general and comprehensive account of not only verbal derivation but also nominal derivation in Hindi. We now proceed to show how this can be done.

While discussing nominal derivation (4.1.2.1), we stated that in Hindi there is free variation between a short mid and the corresponding short high vowel in many cases. This free variation is generally found in those cases where a long mid vowel is shortened because of a derivational suffix being added to the stem. For example, mōṭa 'fat', but both moṭāpā/muṭāpā 'obesity', ēk 'one', but ekkā/ikkā 'ace', chōr 'thief', but churānā/chorānā



'to steal', sēth 'a businessman', but sithānī/sethānī 'businessman's wife', etc. The same situation also exists here. For example both rukṡānā/rokṡānā 'to cause to stop', tukṡānā/toṡṡānā 'to cause to break', chhutānā/chhoṡānā 'to get free', bikṡānā/bechṡānā 'to cause to sell', etc. exist in pronunciation. Of course, it is true that the pronunciation with a short high vowel is preferred in both Hindi and Urdu, at least in the elegant varieties. But this obviously should belong to the stylistic or the sociolinguistic aspect of the situation rather than to a purely phonological one. We think that the pronunciation with the short mid vowels is because of borrowing from Avadhi and other NIA dialects current in the Hindi speaking area. If so, this means that Hindi has an extra rule for raising the short mid vowels after the application of SVS to the stem resulting in the shortening of the long mid vowel. We might formulate this rule as follows.

PR10: Mid Vowel Raising (MV Raising)

$$\left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ -\text{hi} \\ -\text{lo} \\ -\text{long} \end{array} \right] \rightarrow [+hi]$$

Now given this rule, the verbs of the type given in Table 3 can be perfectly well integrated with those given earlier in Tables 1 and 2; in fact the whole paradigm of verbal and nominal derivation can be subsumed under one very general

To account for these verbs we need a rule deleting the a of the diphthongs optionally in the relevant contexts. This will be a minor rule and so the verbs in question will be marked for undergoing it as well as for not undergoing SVS, which is a major rule. If this rule is ordered before the high vowel, which comes after a, is made non-syllabic, the high vowel will remain as the nucleus of its syllable. Alternatively, if we want to avoid this ordering constraint, we can formulate a glide vocalization rule which will turn [-syll] into [+syll] when the glide is between two consonants. Either way there will be no problem with our output. The shortening of the diphthongs can be accounted for in either of two ways. First, we can have a minor rule for reducing a to ə in the appropriate context; there will be an either/or condition imposed on these two minor rules, i.e., the rule deleting a and the rule shortening a, so that only one of them can apply to a given form. Second, it is possible that this reduction rule might not be needed as a separate rule for these few verbs only, and that the reduction could be taken care of by stress shift. This alternative seems to be specially promising as in the forms given above the weakening of the diphthong does not depend on suffixation as such, which is the case elsewhere, but rather on whether the suffixation results in a stress shift or not. For example, there is no weakening of the

diphthongs in lauṭvānā, taulvānā etc., but there is a weakening in lauṭānā, taulānā, baithānā etc. The reason is that in the first case the diphthong is in a closed syllable, and a vowel is always stressed in a closed syllable in Hindi. But in the second case, the stress is not on the diphthong but on the following ā; as such the diphthong is automatically reduced or weakened.

Although it is not our aim to go into the matter of stress in Hindi in any detail, some remarks on the subject will not be remiss here. Syllables in Hindi are stressed if they are either closed or if they contain a long vowel or a diphthong, i.e., if they are strong or heavy syllables (Kelkar 1968:25ff). Normally there is only one main stress in a word. From this it follows that if because of derivational processes two independently stressed syllables come together in a word, some adjustment in the stress pattern is required. If the first syllable contains a long vowel and is open, it loses its stress in favor of the main stress on the following syllable. This is precisely what happens in the case of verbs like lauṭānā 'to return (tr)' etc. But if the first syllable is closed, then it does not lose its stress; this is what happens in the case of verbs like lauṭvānā, taulvānā etc., and consequently in these cases no weakening of the diphthong takes place. Of course, it is possible that there is a weakening of the

stress but it is not sufficient to result in the shortening of the diphthong. In the absence of an instrumental analysis of stress in such words, we are unable to say anything further on the matter.

Coming back to our main theme, we can say that almost the entire verbal derivational paradigm can be explained in terms of suffixation and the consequent shortening of the stem vowel through SVS. There will be only a handful of verbs left unaccounted for by the present analysis. A majority of these have only one form. A few might have two and a few others three. A majority of these like hinhinānā 'to neigh', khatkhatānā 'to knock', bhaRbhaRānā 'to knock', jhilmilānā 'to shimmer', etc. are echo formations. None of these will present any insurmountable difficulties. Thus we can say that in Hindi, both in verbal as well as in nominal derivation, it is the quantitative alternations that are found and not qualitative ones. In fact real qualitative alternation is found in the case of only one verb, i.e., khānā 'to eat', which has khilānā 'to serve food' and khilvānā 'to cause to serve food', though here also khavānā exists as a dialectal variant.

We think that the foregoing treatment of nominal and verbal derivation has given sufficient support for our decision to use the long/short distinction for vowels rather than the tense/lax one as Pray does, for example.

Although it is perfectly possible to formulate SVS in terms of the tense/lax distinction, we feel that a statement of these alternations in terms of vowel quantity is much more intuitively satisfying than one in terms of vowel tenseness. (Also see 1.5.1.) Finally, we would like to state again that we have attempted to present both nominal and verbal derivations as part of a single phonological process, viz., that of stem vowel shortening. In this way we think we have been able to meet, at least partially, the difficulty raised by Pray about the traditional treatment of vowel alternations in Hindi verb stems. Pray writes that "traditional grammars of Hindi-Urdu discuss the vowel alternations of verb-stems but make no attempt to relate these alternations to similar phenomena elsewhere and thereby miss a very important generalization about the phonological structure of tadbhava forms" (Pray 1970:95). By relating the vowel alternations in verb forms to those in nominal forms, we have tried to capture this important generalization in our limited way.

#### 4.1.2.3 Tadbhava compounds

In the preceding two sections we have seen that vowel length is predictable in Hindi in derivational morphology, i.e., when a derivational suffix is added to a stem, whether a noun or a verb or an adjective, the stem vowel is shortened.

There are two more situations in which vowel length is predictable; these are: First, tadbhava, i.e., native, compounding and, second, the plural inflection of nouns ending in long high vowels. We shall discuss these in the order given.

The principle behind the shortening of the stem vowel in compounding seems to be the same as that in the case of derivational morphology. In fact, both can be taken as exemplifying the same general process of word formation, one by the addition of a suffix and the other by the addition of another word. As such both can be related to some general constraint in Hindi regarding the number of moras in a Hindi word. (We take up this question in Chapter 7.) Be that as it may, our concern at present is with the fact of vowel shortening rather than with reasons behind it. Historically, it might have been the case that many formatives that are suffixes now were words in their own right. This would give some explanation for vowel shortening in both the cases; however, synchronically, there is absolutely no motivation for treating the suffixes as independent words. Consequently we have kept the two processes apart.

Consider, for example, compounds like baṭ-mār 'highwayman': bāṭ 'road', bat-kahī 'altercation': bāt 'talk (n)', pāch-mēl 'of five kinds': pāch 'five', ik-tārā 'one-stringed instrument': ēk 'one', kal-mūhā

'black-faced (m,sg)': kālā 'black (m,sg)', tigunā 'three-fold': tīn 'three', pan-dubbī 'submarine': pānī 'water', bhukh-marī 'famine': bhūkh 'hunger', dugunā 'two-fold': dō 'two', etc. We have hyphenated the compounds to separate the two members. It is clear that the shortening of the long vowel of the first member of the compounds is perfectly well predictable. We would like to add that this shortening of vowels is typically an NIA development as Sanskrit did not have any such rule; for example, there is no shortening in the case of tatsama, i.e., Sanskrit, compounds like kāla-mukha 'black-faced', ēka-akṣariya 'monosyllabic', rāja-kula 'royal family', but cf. the native rajvārā 'prince', etc. However, as far as dvandva compounds are concerned, there is no shortening in either tadbhava or tatsama compounds; as examples of tadbhava dvandva compounds we can take mā-bāp 'mother and father', āg-pānī 'fire and water', rājā-rānī 'king and queen', etc.

#### 4.1.2.4 Shortening in Plurals

The shortening of a final long high vowel in plural inflection is the only example of vowel shortening in inflectional morphology. More precisely, the final long high vowel of a noun is shortened when the plural suffix, either direct or oblique, is added to the noun, e.g., nadī/nadiā 'river/rivers', ghōRī/ghōkiā 'mare/mares', bahū/bahuē 'wife/wives', mālī/maliō 'gardener/gardeners (obl,pl)', sādhū/sādhuō 'saint/saints (obl,pl)', ālū/āluō

'potato/potatoes (obl,pl)', etc. As the shortening takes place in the case of all words whether Sanskrit or Hindi, we have not indicated this fact in the above examples. We have also not indicated whether the final long vowel is original or derived through FVL (see 4.1.1.2), as this fact is also immaterial here. It can be said that as FVL applies only when the high vowel is word final, the question of shortening in the case of high vowels subject to FVL does not arise. The moment the plural suffix is added to the noun, the final vowel will cease to be word final, as such FVL cannot apply and the original short high vowel will remain as it is. No doubt the point is well taken and we can account for the short high vowels in forms like sādhuō 'saints (obl,pl)', kaviō 'poets (obl,pl)', etc. on this basis. However, this will not account for the short high vowels in the plurals of many more words with final original long high vowels, both Sanskrit and Hindi, e.g., nadiā 'rivers', bahue 'wives', etc. For such words we shall require a rule which will shorten the high vowel in the relevant context. We might formulate the rule as follows.

PR11: High Vowel Shortening

$$\left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{hi} \end{array} \right] \longrightarrow [-\text{long}] \quad / \quad \text{---} + \begin{array}{c} \text{X} \\ [+pl] \end{array}$$



## 4.2 VOWEL NASALIZATION

### 4.2.1 Background

Nasalsed vowels are found in all NIA languages and on surface they contrast with the oral vowels as well. For this reason they have been taken as phonemic by the Neo-Bloomfieldian linguists working in this area. (See Dixit 1963 and Kelkar 1968 for Hindi, Pandit 1957 for Gujarati, Ferguson and Chowdhury 1960 for Bengali, Kelkar 1958 for Marathi, etc.) As the phonemic theory followed has been of once -a- phoneme-always-a-phoneme type, no effort has been made to examine whether nasalsed vowels have only one source in all cases or more than one; in fact, given the theory no such effort was possible. Once a segment is accepted as a phoneme, there the matter ends. Among generative phonological treatments of vowel nasalisation we can mention Narang and Becker 1971 (=NB), Ohala 1975, Bhatia and Kenstowicz 1972 (=BK). NB have proposed to derive nasalsed vowels from an underlying vowel + nasal consonant sequence. BK have pointed out some ordering problems in NB's treatment and have consequently stated that NB's proposal will not work. They have, in turn, proposed that vowel nasalization is phonemic, i.e., distinctive, in Hindi and cannot be derived through a rule, at least not until some global derivational constraints are placed on the vowel nasalization rule. Thus, they have ultimately held the SPE theory responsible for NB's failure to account for vowel nasalization through their rule.

We take up the question of vowel nasalization in Hindi with two aims in view. First, we would like to examine whether nasalised vowels in Hindi have only one, or more than one, source. Although we are not against the principle of rule optimization, we think that too rigid an adherence to it is undesirable as it can lead to unmotivated analysis. As an example, we can cite Chomsky and Hall's deriving the diphthong in words like kite, i, right, etc., also through the Vowel Shift Rule (see Kiparsky 1968b, Vachek 1970 for a discussion). Second, we would like to examine BK's contention that it is because of a weakness in the SPE theory that NB have not been able to derive nasalised vowels in Hindi through a rule and that, therefore, stronger devices like global constraints on derivations are needed.

#### 4.2.2 Proposed Treatment

We shall consider the question of vowel nasalization under two broad headings. First, nasalization in the presence of a nasal segment, and second, nasalization without the presence of a nasal segment. We shall take these up in the order given.

##### 4.2.2.1 Nasalization in the presence of a nasal segment

As far as the first situation is concerned, matters are really very simple. In fact, the nasalization of a vowel in the presence of a nasal segment is very common in

natural languages. Consider, for example, words like ām 'mango', ān 'prestige', āndar 'inside', mān 'respect', kām 'work', kāṅhā 'comb', āmbar 'sky', āndhā 'blind', dāṇḍā 'stick', gānā 'song', nām 'name', sūndar 'beautiful', mīn 'fish' etc. Clearly the nasalization of the pre-nasal vowels in these words is due to the presence of the following nasal and as such is entirely predictable. We do not have to represent such pre-nasal vowels as [+nasal] in the underlying representations. Instead we need a rule like the following.

PR12: Vowel nasalization rule (VNR)

$$\left[ \begin{array}{c} +\text{voc} \\ -\text{cons} \end{array} \right] \longrightarrow [+nas] \quad / \quad \text{---} \quad [+nas]$$

It may be the case that for certain native speakers both the pre- and post-nasal vowels are nasalized. That is, they might have nasalized vowels in words like nāch 'a dance', nātā 'relationship', mās 'month', māt 'defeat', nal 'faucet', mukh 'face', etc. rather than the more usual oral vowels.

These speakers can be said to have another rule which would be a mirror image of the rule proposed above. These two rules can be combined in one rule following either Bach's Neighborhood Convention or Langacker's Mirror Image Convention (see Bach 1968, Langacker 1969). Expressed as one rule, such a rule would state that a vowel is nasalized when it is either pre- or post-nasal. However, as this rule

would be only for a minority of Hindi speakers, and as the majority of them have nasalization in the pre-nasal context only, we shall retain only the earlier rule, VNR, in our grammar.

#### 4.2.2.2 Nasalization in the absence of nasal segment

Next we pass on to the nasalization of a vowel without the presence of a nasal segment. Here again we can decide our data into two sets. In the first set we get words with related forms such that in some of these forms a nasal segment is present, while in some others it is not. Furthermore, if the nasal segment is present in related forms, we get a short nasalised vowel and if it is absent, we get a long nasalised vowel. More precisely, the alternation is between VNC and  $\bar{V}C$ . In the second set we get words which in all their related forms are without the presence of a nasal. As an example of the first set, consider alternations like ādhī 'duststorm' / āndhī 'blind (f, sg)', kāpnā 'to tremble' / kāmpān 'trembling (n)', kāndhā / kādhā 'shoulder', ḍār 'punishment' / ḍāṇḍā 'stick (m,sg)', kātā / kāntak 'thorn', pāch 'five' / pānch(ānan) 'five (faced)', ānchal / āchal 'saree-front', etc. As examples of the second set, consider words like hāsnā 'to laugh', phāsnā 'to be ensnared', sās 'breath', mūh 'mouth', chōch 'beak', sīchnā 'to irrigate' sichnā 'to be irrigated', phēknā 'to throw', phiknā 'to be thrown', etc. It will be noticed that there are both long

and short vowels as well as some verbs which are morphologically related in the above list. But there is no nasal segment present in any of the cases. We should perhaps mention that in certain cases there exist Sanskrit words with an overt nasal segment from which the Hindi words in question have been derived. For example, sīnchan 'irrigation' from which sīchnā 'to irrigate' is derived, chānchu 'beak' from which chōch is derived. However, as these Sanskrit words are used only in a scholarly style, we have disregarded them. And even if we include these Sanskrit words, there will still be many Hindi words belonging to the second set in whose case nasalization cannot be explained on the basis of their cognates in Sanskrit. For example, Skt. śvāsa > H sās 'breath', Skt. hasta > H. hāth, 'hand', Skt. has > H. hās 'laugh', Skt. aśru > H. āsū 'a tear', Skt. sarpa > H. sāp 'snake', etc. Hence, we shall maintain the two-fold division of the data that we have given above.

#### 4.2.2.3 Solution

The nasalised vowel in the words belonging to the first set will present no problem in derivation; the underlying form will have VNC, and the nasalisation of the vowel will be taken care of by VNR. This will give us one of the surface forms. For getting the form with a long nasalised vowel, we need a rule of the following type.

TR3: Nasal Deletion Rule (NDR)

$$\begin{array}{ccccccccc}
 & & \begin{bmatrix} +\text{voc} \\ -\text{cons} \end{bmatrix} & & \begin{bmatrix} +\text{nas} \\ +\text{cons} \end{bmatrix} & & [+cons] & + & & & & & & & & \\
 & & 1 & 2 & & 3 & 4 & 5 & \rightarrow & 1 & 2 & 3 & 4 & 5 \\
 & & & & & & & & & & [+long] & \emptyset & & & 
 \end{array}$$

As VNR will be ordered before NDR, there will be no problem; the vowel will have been nasalised before the nasal is deleted.

However, it is clear that NDR is simply a variant of CVL (4.1.1.1), which we reproduce below for convenience.

TR2: CVL

$$\begin{array}{ccccccc}
 \begin{bmatrix} +\text{voc} \\ -\text{cons} \end{bmatrix} & [+cons] & [+cons] & + & \rightarrow & 1 & 2 & 3 & 4 \\
 & 1 & 2 & 3 & 4 & [+long] & \emptyset & & 
 \end{array}$$

2 subsumes both a non-nasal as well as a nasal consonant. It is also possible to combine CVL, VNR and NDR into one rule. For example, consider the following rule.

TR2': CVL(R)

$$\begin{array}{ccccccc}
 \begin{bmatrix} +\text{voc} \\ -\text{cons} \end{bmatrix} & [+cons] & [+cons] & + & \rightarrow & 1 & 2 & 3 & 4 \\
 & 1 & 2 & 3 & 4 & [+long] & \emptyset & & \\
 & & & & & \langle +\text{nas} \rangle & & & 
 \end{array}$$

This rule says that if the deleted segment is a nasal, then the preceding vowel will be nasalised as well. But this rule

will be able to do away with VNR only in those cases which are subject to CVL(R) as well. It will be unable to give the correct output in those cases where the nasal is not deleted. Thus, it will not be able to give the correct output in the case of a word like āṇḍar 'inside' as the nasal is not deleted in this case. Also, in the case of kāṇdhā / kādhā 'shoulder' it will give the correct output only in the case of the latter as the nasal has been deleted there but not in the case of the former where it has not been deleted. Thus, it is clear that even with the CVL(R) we shall still require VNR for many words. As there is hardly anything to be gained by this treatment, we reject it in favor of our earlier analysis and retain both VNR and CVL in this order, NDR being unnecessary.

There is a similar problem here as we had earlier (4.1.1.1) while discussing CVL in connection with quantitative vowel alternations. That is, in some cases the rule seems to be obligatory while in some others it seems to be optional. The explanation is also the same. The rule is obligatory when it is a derivational suffix that is added to the stem, optional when it is an inflectional one. For example, ādhī 'duststorm', ḍāṛ 'punishment' are derived from the respective stems andh- 'blind' and dand- 'stick' after the addition of derivational suffixes ī and a respectively. (The final a later gets deleted; see Chapter 6

for a discussion.) On the other hand, when simply an inflectional suffix is added, e.g., the masculine singular marker ā, we can optionally get either kāṇdhā or kādhā 'shoulder'. Again, we think that in the case of doublets, the forms with long nasalised vowels are borrowed from Avadhi and other NIA dialects current in the Hindi speaking area. Thus, if we are correct, the free variation between such forms is a result of dialect interference.

While discussing CVL in connection with quantitative vowel alternations (see 4.1.1.1) we said that the justification for deleting the first rather than the second consonant after the vowel will be given while discussing vowel nasalisation. Here we find that it is the nasal, i.e., the first post-vocalic consonant, which is deleted. It would be much simpler if the consonant occupying the same position is deleted in all cases. A further justification for the deletion of the first rather than the second consonant can be found in those cases also where no nasal segment is involved. Consider, for example, alternations like matthā / māthā 'forehead', addhā / ādhā 'half', matthā / māthā 'buttermilk', etc. In these cases also we find that it is the unaspirated, i.e., the first, consonant which is deleted.

4.2.3 What we have done so far enables us to derive nasalised vowels in all cases except one, viz., the second



set of our data given above. This set consists of those words where no nasal ever surfaces. Consider, for example, words like ĩṭ 'brick', āsū 'tear', bēchnā 'to sell', hāsnā 'to laugh', hāsī 'laughter', phēknā 'to throw', phiknā 'to be thrown', ākḥ 'eye', sās 'breath', sīchnā 'to irrigate', sichāī 'irrigation', etc. Some of these words illustrate the phenomenon of spontaneous nasalization in MIA from where they have come into NIA. That is, the OIA cognates of these words did not have a nasal in them, e.g. āsū <

Skt. aśru 'tear', ākḥ < Skt akṣin 'eye', hās < Skt has 'to laugh', sās < Skt śvāsa 'breath', etc. However, not all words are of this type. More important, we are concerned with the language as it is now, and a historical explanation, through it may help us in understanding a problem, may not necessarily help us in giving a solution to it.

Two treatments are possible to account for the nasalization of the vowel in such words: we might call them the abstract and the concrete treatments respectively. Under the abstract analysis we represent such words as having a nasal segment underlyingly, which is then obligatorily deleted by CVL; VNR applying before CVL will take care of the nasalization. Thus, āsū 'tear (n)' will be represented as ansū, from which VNR and CVL, applying in this order, will give us āsū. However, this treatment will work only in those cases where the nasalized vowel is long. It will

not work in those cases where the nasalised vowel is short; for example, if hās 'to laugh' is represented as hans, VNR and CVL will give us \*hās, a non-existent form. And there are many such words in the language; e.g., khichnā 'to be pulled', phiknā 'to be thrown', siknā 'to be roasted', phāsnā 'to be ensnared', hāsī 'laughter', sikāī 'fomentation', sichāī 'irrigation', etc.

The abstract analysis at this point has no other alternative but to mark all such words as exceptions to CVL. At the same time it has to write another rule, a variant of CVL, which will delete the nasal from these words without lengthening the vowel. In fairness to the abstract analysis it must be said that the number of words requiring this special treatment can be lessened. Thus, verbs like phiknā 'to be thrown', phāsnā 'to be ensnared' and nouns like sichāī 'irrigation' will be derived from phēknā 'to throw', phāsnā 'to ensnare', sich 'irrigate' respectively. As in these cases the vowel is long there will be no problem, and once the vowel has been nasalised and the nasal deleted, in any subsequent derivations also there will be no problem. But there will still remain some words for which no such source is available, e.g., hāsnā 'to laugh', hāsī 'laughter', sābhalnā 'to take care', sāvarnā 'to decorate', etc. And these words will have to undergo the special treatment.

The concrete analysis, on the other hand, will mark all the nasalised vowels in such words as [+nas] underlyingly. It is not the case that this approach has nothing to commend itself; although as the theory of generative phonology stands at present and as it has been applied, it would seem so. Let us examine carefully what the abstract analysis entails. First, by positing a nasal segment in the underlying representation of such words and then obligatorily deleting it, we have made use of something like absolute neutralization. What is more, we have done so without providing any evidence whatsoever in its support. We have used this qualification advisedly as our case is not at a par with the classic cases of absolute neutralization like Chomsky and Halle's positing /x œ/ and the epsilon glide for English, Hyman's positing /e o/ for Nupe, Brame's positing /S/ for Maltese Arabic. (See Chomsky and Halle 1968, Hyman 1970, Brame 1972. For a criticism of this approach see Kiparsky 1968b, 1971, Vachek 1970, Shibatani 1971, Wang 1971.) Kiparsky 1971 has allowed absolute neutralization provided the posited abstract segment has a role to play elsewhere in the grammar too. This is not the case here; as far as we know the posited nasal segment is not required anywhere else in the grammar. Second, we have had to add a minor rule of nasal deletion for the ḥasnā-class of words. Note that

this rule is contrary to the general tendency of compensatory vowel lengthening not only in Hindi but also in NIA, a tendency which is admirably captured by the CVL. Note further that these words are purely native and hence we cannot appeal to coexistent phonemic systems. (See Fries and Pike 1949.) Third, we have had to mark such words as undergoing this minor rule. Fourth, we have had to mark such words as not undergoing CVL, the major rule. Lexical marking, we may add, is a very costly affair in a grammar and should not be resorted to unless unavoidable. Also, it makes a form idiosyncratic. It might be said in reply that these forms are irregular or idiosyncratic and the lexical marking and a new rule capture just this irregularity. However, it is not clear to us in exactly what sense these words can be said to be irregular. It is simply the presence of a nasalised vowel without the presence of a nasal segment which can be called irregular, only in the sense that the nasalised vowel cannot be derived by a rule which covers so many cases in the language. But by representing the vowels in question as [+nas] in the underlying representation, it is precisely this irregularity that we shall be capturing without implying anything about the forms as a whole. Not only this. We shall also be saving ourselves one minor rule and two lexical markings. Even for an overtly formalistic and rule-dominated approach this should

be a desirable aim. No doubt the optimization of a rule is desirable; so also is the derivation of a particular segment in the same way in all cases as far as possible. But we do not believe in the once-a-derived-segment-always-a-derived-segment approach; it is too much like the old phonemics. (In fact, as is clear, even the optimization of an existing rule is not achieved by the abstract analysis.)

On the basis of this discussion we feel that the concrete analysis is much simpler than, and hence superior to, the abstract one; consequently we propose to adopt it. In other words, we are claiming that in Hindi phonetically nasalised vowels can come from two sources. There are some nasalised vowels which can be derived by VNR and, thus, are not marked [+nas] in the underlying representation. There are, however, some other nasalised vowels which cannot be derived from VNR and are marked [+nas] in the underlying representation. Needless to say, it is these nasalised vowels alone which can be really called phonemic, that is, if we care to use the term.

#### 4.2.4 Conclusion

We stated in 4.2.1 that we had two aims in undertaking the study of vowel nasalization in Hindi. The first was to examine whether nasalised vowels had only one,

or more than one, source. So far, nasalised vowels have been attributed to only one source. This is true of traditional phonemic analyses as nasalised vowels were taken to be phonemic; but we also find the same view expressed in generative phonological studies of Hindi, e.g., NB derive nasalised vowels only by a rule; in contrast, BK and Ohala 1975 prefer an underlying [+nas] specification. Ohala 1975 criticises NB for setting up abstract underlying representations like /dhuāN/, /bhauNrā/ for words like [dhuā̃] 'smoke (n)' and [bhāṽrā] 'black bee' respectively and deriving the latter from the former by ad hoc phonological rules. The representations are abstract in the sense that the posited nasal segment never surfaces phonetically. So far we agree with her. However, she seems to be of the opinion that as all the nasalised vowels cannot be derived through a rule like  $V \rightarrow V/-N$ , it is pointless to have a vowel nasalization rule of the above type. Again, it is an all-or-nothing approach, an approach that we fail to appreciate. We do not think that the theory anywhere states this as a basic requirement for doing generative phonology. In fact, had this been the case, linguists like Kiparsky would not have been recognised as generative phonologists, which is certainly not the case.

We must state again that we are not against the optimization of a rule; in fact, we have ourselves taken

that approach in the case of aspiration in Hindi (see 2.1), where we derive all aspirated consonants from an underlying Ch sequence. But there, as we have tried to show, our analysis is able to explain certain synchronic and diachronic facts about not only Hindi but also OIA and MIA. These facts cannot be explained if we take the unit segment approach for aspirates within a formative and the cluster approach for aspirates when a word final unaspirated consonant is followed by h. Note that this dual approach, which we rejected in the case of aspirates is, in principle, exactly similar to the one we have taken in the case of nasalised vowels, viz., positing two different sources for the same segment. But the point is that in the case of nasalised vowels no maximization of the explanation provided is possible as a result of the rule optimization approach, as was the case with aspirates. Considering the fact that Modern Hindi, apart from having *tadbhava* and *dēshaj*, i.e., non-Sanskrit, forms, has a large number of borrowings from Sanskrit and Arabo-Persian, which are languages with widely differing phonological systems and constraints, it would indeed be surprising if a single rule could account for all the facts of a given surface type in the language. The diversified nature of underlying sources for what are superficially similar phenomena is precisely what we find reflected in our dualistic approach to nasalized vowels in Hindi.

Our second aim in taking up the question of vowel nasalization was to examine whether there was any inherent weakness in the SPE theory which prevented NB from accounting for nasalised vowels by a rule, and, therefore, whether devices like derivational constraints were needed in the grammar. We have shown that the weakness does not belong to the theory but rather to the particular way the theory has been applied. Consequently, once we reject this interpretation, we need not have any problem with nasalised vowels in Hindi. Needless to say, that as far as these particular facts from Hindi are concerned, they provide absolutely no evidence in favor of incorporating derivational constraints in the theory.



## CHAPTER 5

### ON MID VOWELS

#### 5.1 ALTERNATIONS INVOLVING MID VOWELS

In the earlier chapters we did not say anything about the underlying status of mid vowels in Hindi; specifically we did not discuss whether they occur as such in underlying representations, giving us a five-vowel system /a i u e o/ for Hindi, or whether they are derived from some other vowels, or sequences of vowels, giving us probably a three-vowel system /a i u/. For example, while discussing vowel alternations in the verbal derivational paradigm (4.1.2.2), we stated that in the case of alternations like biknā 'to be sold' / bēchnā 'to sell' / bikvānā 'to cause to sell', rukṇā 'to stop (intr)' / rōkṇā 'to stop (tr)' / rukvānā 'to cause to stop', etc. the forms with the mid vowels are to be taken as base forms from which the other forms can be derived by the addition of a derivational suffix and the application of SVS (4.1.2.1). Thus we tacitly assumed that the mid vowels in such cases occur in the underlying representations, and presumably elsewhere too. In this chapter we would like to examine this question in some detail.

## 5.2 THE HINDI VOWEL SYSTEM: THE DERIVATION OF MID VOWELS

Examining the question of the representation of mid vowels is important from two points of view. First, if the mid vowels can be derived from some other source, we can get a simpler underlying vowel system. The vowels or the sequences of vowels from which these mid vowels can be derived can only be from the set /a i u/, as these are the only other vowels which can occur at the underlying level. Thus our underlying vowel system will consist of only /a i u/, the optimal vowel system according to the SPE (p. 409). Second, and it is related to the first, it is agreed that the underlying vowel system of OIA was precisely this. There have of course been many changes between OIA and NIA, but if the underlying vowel system has not changed, then we can say that, at least as far as Hindi is concerned, there has been no restructuring in this respect and the changes are due solely to a change in the rule component. This will have obvious implications for IA studies. For example, it might be possible to show that all NIA languages have the same underlying vowel system as OIA. If so, it would be interesting to examine the different rules in OIA and NIA for deriving the mid vowels from the same underlying representations and see in what respects, if any, they differ from each other and whether any simplification has

taken place or not. In other words, it would have importance for both descriptive and historical studies of IA.

Let us start with the hypothesis that Hindi has only /a i u/ at the underlying level and that mid vowels are derived from them. We can examine this under two headings. First, their derivation from single vowels, and second, their derivation from a sequence of vowels. We shall take these up in the order given.

### 5.2.1 Derivation from Single Vowels

In view of the alternation of ē with i and of ō with u, for which there is ample evidence in both the nominal and verbal paradigms (4.1.2.1, 4.1.2.2), we can posit i, u in the underlying representations and derive ē, ō respectively from them by a process of lowering and lengthening. As far as the free variation between the short mid and the short high vowels is concerned, it can be taken care of by the Mid Vowel Raising rule (PR10, 4.1.2.2), a rule which is of fairly general application in Hindi. It is to be noted that this rule is required whether we derive the mid vowels from high vowels or posit them in the underlying representations as such. Before we proceed any further, we would like to state the various alternations that the mid vowels have; in this connection we shall consider both the verbal as well as the nominal paradigms.

### 5.2.1.1 Types of Alternations

These alternations can be divided into the following three groups.

1. Group 1 - The long mid vowel alternates with its short high counterpart. For example, bēchnā 'to sell' / biknā 'to be sold' / bikvānā 'to cause to sell', rōknā 'to stop (tr)' / ruknā 'to stop (intr)' / rukvānā 'to cause to stop', phēknā 'to throw' / phiknā 'to be thrown' / phikvānā 'to cause to throw', khōlnā 'to open (tr)' / khulnā 'to open (intr)' / khulvānā 'to cause to open', etc.

2. Group 2 - A long mid vowel, more precisely, a long back mid vowel, alternates with a long high back vowel. For example, tōknā 'to break (tr)' / tūtnā 'to break (intr)', chhōrnā 'to leave (tr)' / chhūtnā 'to leave (intr)', phōrnā 'to burst (tr)' / phūtnā 'to burst (intr)', sōkhnā 'to absorb' / sūkhnā 'to dry (intr)'.

3. Group 3 - A long mid vowel alternates with either a short mid vowel or its short high counterpart. For example, khētī 'farm' / khetihar 'farmer', ēk 'one' / ikkā/ekkā 'ace', sēk 'foment' / sīkāi/sēkāi 'fomentation', khōd 'dig' / khudāi/khodāi 'digging', mōtā 'fat' / mutāpā/motāpā 'obesity', sōnā 'gold' / sunār/sonār 'goldsmith', khēl 'play (n)' / khilārī 'player', ghōRā 'horse' / ghuRdauR /

ghoṛdauR 'horse race', mēlā 'fair (n)' / milnā 'to meet', etc. This is by far the largest group.

Of these three groups, Group 2 is clearly an exception and should be treated as such. As far as Groups 1 and 3 are concerned, they can be subsumed under one larger group; regarding this group we can say that a long mid vowel alternates with either its short counterpart or with its short high counterpart. In some cases only one of the two possible alternants is found, whereas in some other cases both are found. We might add that the formal or correct styles of both Hindi and Urdu **prefer** the short high vowel and that the forms with short mid vowels are due to dialectal borrowing. It is possible that because of this reason, in our Group 1, the variants with the short mid vowel have been totally eliminated.

### 5.2.2 Accounting in the Verbal Paradigm

Thus we can represent our long mid vowels as underlying short high vowels, i.e., i for ē and u for ō. As far as the verbal paradigm is concerned, it will make the more frequent variant the underlying vowel. Note that it is only the class II, i.e., base, forms like rōknā 'to stop (tr)', khōlnā 'to open (tr)', bēchnā 'to sell', khēlnā 'to play', etc., that we get the long mid vowels. In the rest of the related forms like rukṇā 'to stop (intr)' /

rukvanā 'to cause to stop', biknā 'to be sold' / bikvanā 'to cause to sell', khilvanā 'to cause to play' / khilārī 'player', etc., we get the short high vowel. The following rule will take care of the class II or base forms.

PR13: Class II Rule

$$\left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{hi} \\ -\text{long} \end{array} \right] \longrightarrow \left[ \begin{array}{l} -\text{hi} \\ +\text{long} \end{array} \right] / \text{X} - \text{Y} \quad [\text{Cl II}]$$

For the present we shall assume that the free variation between short mid and short high vowels can be in either direction, i.e., mid to high as well as high to mid; consequently, we formulate the following rule to take care of this variation. Note that Mid Vowel Raising (PR10, 4.1.2.2) is unidirectional, i.e., mid to high only. We should mention that we shall eventually generalise our account and shall find that Flip as now formulated is unnecessary and its purpose can be served by constraining Mid Vowel Raising Rule (PR10). (See 5.2.5 below.)

PR14: High-Mid Flip (Flip)

$$\left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ \alpha\text{hi} \\ -\text{lo} \\ -\text{long} \end{array} \right] \longrightarrow / \quad [-\alpha\text{hi}]$$

As the class II Rule will apply in the case of Class II, or base, verbs alone, it is only here that we shall get a long mid vowel; in the other forms the short high vowel will

remain as it is. Flip may apply optionally in some cases, obligatorily in some others, and not at all in still others; this can be taken care of through marking.

### 5.2.3 Accounting in the Nominal Paradigm

Let us see if we can extend our analysis to the nominal paradigm as well. Here we shall be mainly concerned with those nouns with long mid vowels which cannot be derived from a corresponding verbal root which shows a long high vowel. That is, these nouns exist as basic nominal stems rather than as derived ones.

Consider, for example, alternations like khētī 'farm' / khetiḥar 'farmer', ghōRā 'horse' / ghuRdauR 'horse race', sōnā 'gold' / sunār 'goldsmith', lōhā 'iron' / luhār 'ironsmith', thēlā 'cart' / thiliā 'cart (dim.)', ēk 'one' / ikkā 'ace', mōtā 'fat' / mutāpā 'obesity', etc. The first member of each pair is either a noun or an adjective; the second, which is clearly derived from the first by the addition of a suffix, is always a noun. Thus, the stem or the base form can be represented as having an underlying short high vowel; there is a rule which changes this vowel to the corresponding long mid vowel in the case of stems; in the case of derivation the rule does not apply, and hence we get the short high vowel. Its free variation with the short mid vowel will be taken care of by the High-Mid Flip.

Thus, the output will be correct in all cases.

However, it is clear that in the case of nouns we have done the same thing which we did in the case of verbs. There too we had underlying short high vowels raised to long mid vowels in the class II form of verbs by the Class II Rule; here we require the same thing to be done for nominal stems. We can combine the two rules into one as follows.

PR15: High Vowel Lowering (HVL)

$$\left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{hi} \\ -\text{long} \end{array} \right] \longrightarrow \left[ \begin{array}{l} -\text{hi} \\ +\text{long} \end{array} \right] / \quad X \text{ --- } Y \left\{ \begin{array}{l} [+Cl \text{ II}] \\ \text{S} \\ [+stem] \end{array} \right\}$$

Here S stands for substantive; this way both nouns and adjectives have been taken care of. As far as the free variation between short mid and short high vowels is concerned, it will be taken care of by Flip in both cases.

#### 5.2.4 Exceptions

Although HVL has worked so far and we have been able to account for alternations in both the verbal and nominal paradigms, it will not work in all cases. As the rule changes short high vowels into long mid ones in the base



form of verbs and in nominal stems, it follows that in these two forms we should never get a short high vowel, as all instances of it will be changed to long mid by HVL. Unfortunately, this is not the case. Although in a number of cases the short high vowel can be derived from the corresponding long one, i.e., underlyingly we have a long high vowel, yet there remain a large number of cases where this is not possible. Wherever the short high vowel is derived from the long one by SVS, there will be no problem; we simply have to order HVL before SVS. However, where the short high vowel cannot be derived from the long high vowel by SVS, i.e., in those cases where we get a short high vowel in all the forms and, hence, the short high vowel has to be taken as underlying, there will be a problem. What is more, there are many such cases both in the verbal as well as in the nominal paradigms. For example, consider forms like nikalnā 'to come out' / nikalvānā 'to cause to come out', ghisnā 'to rub' / ghisvānā 'to cause to rub', phisalnā 'to slide' / phislānā 'to cause to slide', ulatnā 'to overturn' / ultānā 'to cause to overturn', bichhnā 'to be spread' / bichhānā 'to spread' / bichhvānā 'to causeto be spread', nominal stems like imlī 'tamarind', guR 'jaggery', inām 'prize', bijlī 'electricity', ulāhanā 'complaint', kilā 'fort', pur 'city', khur 'hoof', sir 'head (n)', subeh 'morning', and adjectives like ḍublā 'lean (m,sg)',

sūndar 'beautiful', suhānā 'pleasant (m, sg)', etc. None of these can be accounted for as none of these can be derived from any other formative having a long high vowel.

Another set of words not amenable to the proposed analysis will consist of those in which both a short high vowel and a long mid vowel are found. Again, there are many such words; consider, for example, kubēr 'god of wealth', purōhit 'priest', nissandēh 'without doubt', vidēś 'foreign country', kubēlā 'bad time', nimēs 'moment', viślēṣan 'analysis', subōdh 'easily understandable', vivēk 'wisdom', and many others like these.

In defence we can say that as far as the first group is concerned, it consists of exceptions to HVL; as such the short high vowel remains unchanged. As far as the second group is concerned, it consists mostly of Sanskrit words; we can again make them exceptions to HVL.

Although by resorting to marking, we have been able to save the analysis, we feel that this has been quite costly. Marking, as we have said earlier also, is a very costly affair and should be resorted to only when all other alternatives fail. This cannot be said to be the case here as we have still one other alternative left, i.e., the derivation of mid vowels from sequences of vowels. Then, marking means classifying a particular word, or a set of

words, as idiosyncratic, and we do not know in what sense these words can be called idiosyncratic. It is true that the second group consist of Sanskrit words, but it is equally true that many of them are fairly common in Hindi. Because of these reasons, the use of marking in this case seems to us unjustified. We think that even if we had no further criticism of this analysis, what we have said above would have been justification enough for rejecting it. In the following few paragraphs we discuss another criticism of this analysis.

While discussing quantitative vowel alternations in the verbal paradigm (4.1.2.2) we took the Class II form of verbs as the base from which the other three forms are derived by suffixation and the application of SVS. That is, starting from the base rōk-nā 'to stop', we derived ruk-nā, ruk-ā-nā, and ruk-vā-nā from it by the addition of  $\emptyset$ , -ā, and -vā respectively, and the application of SVS and Mid-Vowel Raising (PR10). Note that the verbs that concern us here all have the long mid vowel in the base form. But according to the approach taken here, the base form is ruk-nā from which we derive rōk-nā by the application of HVL. In other words while in the earlier analysis we derived the non-basic, i.e., Classes I, III and IV, forms by a rule, here we derive the basic, or Class I, form by a rule. Thus

there is a clear inconsistency between the two approaches. We can resolve this inconsistency in two ways.

First, we keep the base form as ruk- and we let it undergo HVL obligatorily; this would give us rōk-. Now if the verb is Class II, or active as most of the verbs of Class II are active, there is no problem as the form will remain as rōk. On the other hand, if the verb is non-active, rōk- will be followed by the suffixes  $\emptyset$ , -ā or -vā as the case may be. But now the SVS and after it Flip can apply and give us ruk- followed by the relevant suffixes of course. Although there will be no problem with the output, we have two criticisms to make of this treatment. First, the Class II forms are the unmarked ones and thus should be taken as basic. All verbs have this form whether they have the other forms or not. Consequently, it seems strange to us to derive the unmarked form by a rule, a rule which seems to exist only for these forms. (We disregard nominal stems for the present.) The second, and much more damaging, criticism of this analysis is that it is highly contrived and unmotivated. For example, what is the function of the underlying short high vowels in these verbs? They seem to exist only to become long mid vowels, which in their turn could undergo SVS, which in its turn could shorten them so that Flip could again make them high. Put this way the

whole process sounds absurd, but this is precisely what will happen in the three derived forms. It is also clear that there is no independent motivation for either HVL or the underlying short high vowels in these verbs. But in TGG it is always desirable to have independent motivation for either a rule or an underlying representation.

Second, we can put these verbs in a separate group, a group which exhibits qualitative vowel alternations; that is, we revise our earlier position (4.1.2.2) that Hindi verbs show only quantitative vowel alternations. This group will consist of only those verbs which show a long mid vowel in the Class II form; and will be marked [-SVS]. The underlying representation will have short high vowels, e.g., ruk-. In the Class II form HVL will apply to the base form and give us the form with the long mid vowel, for example, rōk-; in the rest of the cases the underlying short high vowel will surface as it is. The suffixes for the derivational forms remain the same for all the verbs in the language. But this way the domain of SVS will be restricted, in consequence, the comprehensive picture that we built up in Chapter 4 regarding both the nominal and verbal derivational paradigms will suffer.

Thus it is clear that neither of these two ways is entirely satisfactory. If we follow the first, we preserve

the comprehensive character of SVS but are left with a highly unmotivated rule (HVL) and underlying representations. On the other hand, if we accept the second, although the unmotivated character of the rule and the underlying representations is mitigated, we lose the comprehensive character of SVS and consequently that of the entire nominal and verbal derivational paradigms. We ignored the nominal paradigm deliberately as we felt that sufficient evidence has been presented to enable us to reject the proposed analysis of deriving long mid vowels from underlying short high vowels. We now turn to an examination of the second alternative, i.e., the derivation of long mid vowels from a vowel sequence.

#### 5.2.5 Deriving Mid Vowels from Vowel Sequences: An Alternative Approach

The only vowel sequence from which a mid vowel can be conveniently derived is that of a low back vowel followed by a high vowel; i.e., ai for ē and au for ō. However, as we have shown the feasibility of the earlier analysis, no matter how costly, our task here will be much more difficult. In the first case our aim was simply to show that the analysis will work. Here, as it is an alternative, we shall have to show not only that it is feasible but also that it is better than the first one, if we want that this analysis

thus leaving only the short high vowel in those contexts. Or, second, we can get the long mid vowel first by MVR and then in certain contexts let SVS and Flip apply in this order; this will also give us the short high vowel in those contexts. Let us now examine these two at some length.

If we follow the first alternative, then given alternations like biknā 'to be sold' / bēchna 'to sell' / bikvānā 'to cause to sell', rukṇā 'to stop (intr)' / rōknā 'to stop (tr)' / rukṇā 'to cause to stop', etc., the underlying forms can be represented as /baik/ and /rauk/ respectively. (We disregard the consonantal alternation in the first case as it is not relevant here.) For Class II forms, MVR will apply giving us bēch and rōk; for the rest of the forms the other rule deleting the a will apply and we shall get bik- and ruk-. As the mid vowels are found only in class II forms, i.e., our base forms, MVR can be constrained to apply only in these cases; in the rest of the forms, i.e., suffixal, the other rule deleting the a will apply giving us the forms with short high vowels. We retain Flip as it is given above. As far as nominal alternations are concerned, for example, khētī 'agriculture' / khētihar 'farmer', mōtā 'fat' / mutāpā 'obesity', ghōRā 'horse' / ghuRdauR 'horse race', thēlā 'cart' / thiliā 'cart (dim)', ēk 'one' / ikkā 'ace', etc., we represent the

base or stem forms with underlying sequences. We let MVR apply when the stems surface as such; we let the rule deleting the a apply when the stem forms are followed by a derivational suffix. Without difficulty the two rules can be adequately constrained so as to apply only in the relevant contexts in the case of both nouns and verbs.

On the other hand, if we accept the second alternative, we let MVR apply first giving us the long mid vowel; this will be the class II form, which is non-suffixal. The other forms are all suffixal, and when the proper suffix is added, SVS will apply and shorten the long mid vowel; after this Flip will apply and give us the short high vowel. Similarly, for the nominal paradigm, where also the non-suffixal forms show the long mid vowel while the suffixal ones show the short high vowel.

Even a very superficial examination will tell us that the first approach leaves a lot to be desired. First, this approach drastically restricts the application of SVS; this in itself may not be very bad. What, however, is definitely undesirable is the fact that the rule deleting the a in certain contexts is totally unnecessary. Note that this rule will apply only in those cases where a derivational suffix has been added. But this is precisely the context for SVS, and we have seen that SVS along with



Flip is perfectly capable of giving us the correct output in all the cases which have been given here as well as many others. Of course, Flip will still be needed, but this rule, or some variation of it, will be needed anyway, whether we have the mid vowels underlyingly or whether we derive them from some other source.

To sum up the discussion so far we can say that if the long mid vowels are represented underlyingly as vowel sequences, i.e., ai for ē and au for ō, we need only one extra rule, MVR. Given MVR and SVS (4.1.2.1), we can take care of all alternations found in the nominal as well as in the verbal paradigms in Hindi. Also, now as we have to account for the free variation between the short mid and the short high vowels in one direction only, i.e., from mid to high, Flip is unnecessary. Note that the free variation is found in the derived forms only, whether nouns or verbs. As such all these forms will be subject to SVS, which will shorten the long mid vowels of the base forms when derivational suffixes are added to the stem or base. We think that the Mid Vowel Raising Rule (PR10, 4.1.2.2) reformulated as a context-sensitive rule will be sufficient for our purposes. Consequently, we reformulate it as follows.

PR10': Mid Vowel Raising Rule (R)

$$\left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ -\text{hi} \\ -\text{lo} \\ -\text{long} \end{array} \right] \rightarrow [+hi] \quad / \quad X \text{ --- } Y + Z \quad \begin{array}{l} [+der \\ +suff] \end{array}$$

It will be naturally ordered after SVS; the reference to a derivational suffix is essential to keep it from applying to short e coming from a (PR1, 2.2.1) and from i (PR3, 2.2.3).

### 5.2.6 Choosing Between Two Alternatives

Thus as far as the feasibility of this analysis is concerned, we have shown it, both in the case of verbs and in the case of nouns. However, as we said in the beginning, because it is an alternative analysis, it would be better if we could find some plus points in this analysis which are absent from the earlier one. If we are able to do so, the choice between the two proposed solutions will no longer remain arbitrary. In fact, even as it is, the analysis is better than the earlier one in the sense that it does not have to resort to such large scale lexical marking as is necessary in the earlier one. That is, this analysis results in a less costly grammar than the previous one, all other things being equal. In what follows, we propose to show that on the basis of historical considerations and the dialect situation in NIA, this analysis scores

over the former. After showing this, we shall return to synchronic considerations.

#### 5.2.6.1 Historical considerations

OIA has two diphthongs, ai and au, or more precisely, āi and āu. The early Vedic vriddhi diphthongs āi and āu were opposed to the guna diphthongs ai and au. In later Vedic, ai and au became ē and ō respectively. (That is why in Sanskrit ē and ō are always long, i.e., two moras.) In Avestan these are preserved as diphthongs, e.g., Av. daeva: Skt dēva 'god', Av. haoma: Skt sōma 'soma', etc. After this monophthongization, as there was no short a left to contrast with long ā in the diphthongs, āi and āu became ai and au respectively. (We are talking here of pronunciation; underlyingly they are still treated as āi and āu by Sanskrit grammar.) This is the position in later Vedic and Classical Sanskrit.

When we pass on to the Middle Indo Aryan (MIA) stage, we find that the Sanskrit diphthongs have met the same fate here as the early Vedic diphthongs met in later Vedic, i.e., ai has become ē and au has become ō, e.g., Skt taila > MIA tēla 'oil', Skt chaura > MIA chōra 'thief' etc. However, MIA has another source for mid vowels as well; Sanskrit high vowels followed by two consonants became, in

certain cases, in MIA short mid vowels followed by, of course, two consonants. Later on, as many such sequences underwent cluster simplification and compensatory vowel lengthening, the short mid vowels became long. For example, Skt sindūra > MIA sendura > NIA sēdur 'vermilion', Skt. udra 'water' > MIA odda > NIA ōdā 'wet', Skt vilva > MIA bella > NIA bēl 'a particular tree', Skt chhidra > MIA chhedda > NIA chhēd 'hole', Skt śimba > MIA semma > NIA sēm 'beans', etc.

These well-attested facts from the history of OIA and MIA have been given simply to show that deriving mid vowels from a sequence of a and the relevant high vowel in Hindi will not be anything new. On the contrary, it will be a continuation of an old process. But we do not claim that the synchronic grammar of a language must recapitulate facts about its history. However, we do claim that if a synchronic grammar, without doing violence to synchronic facts, does recapitulate such facts, it is a much better grammar than one which does not do so, all other things being equal. And it cannot be said regarding this analysis that it fails to account for synchronic facts of the language. In fact, as we have shown above, this analysis is better than the one proposed earlier on purely formal grounds and without any consideration of historical data. After this we pass on to the dialect situation in NIA.

### 5.2.6.2 Synchronic considerations: Dialect situation in NIA

In roughly the west-central part of India, i.e., in and around the Malva region of Madhya Pradesh, four major NIA languages are spoken; these are Marathi, Gujarati, Rajasthani and Hindi. All these four languages have the diphthongs ai and au. The interesting fact is that in Hindi and Rajasthani (Malvi dialect) as spoken here, and we consider the educated speakers of these languages as well, these diphthongs have become monophthongs precisely in the same way as the Sanskrit diphthongs became monophthongs in MIA, i.e., ai has become ē and au has become ō. In Marathi this change has not taken place. As far as Gujarati is concerned, the Saurashtri dialect of Gujarati has undergone this change. In the rest of Gujarati also this change is found, though in restricted contexts (Pandit 1961). We might add that Saurashtra does not border on Malva. In Hindi as spoken here the change has taken place across the board, and has affected all the words with these diphthongs including English loan words. Thus, māi 'I', dauṛ 'run (v)', paisā 'money', aur 'and', ṭaiksi 'taxi', kāinsar 'cancer', kāun 'who', taul 'weight', have become respectively mē, dōṛ, pēsā, ōr, ṭēksi, kēnsar, kōn, tōl, and so on.

The situation in this area is mixed in that both types of pronunciations exist side by side. This is so because

people from the diphthongal area also live here and because the diphthongal pronunciation, being the standard one, is used in radio broadcasts, films, etc. It may be mentioned that the standard pronunciation of all these languages exists outside this region, and the monophthongal pronunciation current here is only a regional one. That is why even the same person may use both the pronunciations depending upon the situation. One remarkable and important fact is that in spite of this widely occurring change, there is no difficulty of communication between the two groups. It should be mentioned that the words containing the diphthongs are not just a handful.

How do we account for these facts? A counter-question can be: Do we have to account for these facts? After all we are not writing a grammar of these languages. It is true that we are not writing a grammar of these languages. But it is equally true that our question is not unrelated to what we are doing here. We are discussing here questions about Hindi phonology, and this is a question about it. Further, we do not think that a grammar of a language should be so narrowly constructed that apart from the specific dialect being studied, it is unable to tell us anything about its other dialects. In fact, one of the greatest dangers inherent in such a narrow interpretation of the term

"specific dialect" or "my/our dialect" is that it might become synonymous with the linguist's idiolect, an unsatisfactory state of affairs on all counts. On the contrary, we believe that the dialects of a language are quite close to each other underlyingly and that their surface differences are generally due to some changes in the rule component (King 1969). As such, dialect variation is always interesting and significant. We are not saying that there is no difference between the grammar of a language and the comparative grammar of languages related to that; we are simply saying that if a synchronic grammar of a language can go beyond the gross facts of the particular dialect of that language under examination and explain facts about related languages and/or dialects as well, then it is a better grammar than one which does not do so, all other things being equal.

#### 5.2.6.3 Problems and Questions in the historical evidence

Let us now turn to another point. We have seen that in late OIA, in all the languages/dialects of MIA, and in three major NIA languages as spoken in and around a particular region, ai and au have become ē and ō respectively. Now why should have this change occurred repeatedly? Why could not ai and au have become æ and ɔ respectively? In WH precisely this change has occurred. (Chomsky and Halle

(1968:360) have this rule for Kasem, an African language, also.) If e and o are derived from underlying i and u respectively, or if they are posited as such at the underlying level, then this widely occurring change repeatedly taking place in Indo Aryan (IA) is an accident. In other words, a change which has taken place in later OIA, in all the MIA languages, and in three NIA languages at least is a chance and as such it is no more significant than a change which, say, would have made all the word-final voiceless stops in all these languages voiced and aspirated. This is too much for us to accept; after all there is a limit to coincidence or chance even in natural languages. But it is precisely this absurd conclusion which a strict separation between synchrony and diachrony or between a language and other related languages leads us to.

On the other hand, if we derive the mid vowels from underlying ai and au, this accident can be very easily and naturally explained. We require only two rules. The first is our MVR, though its precise formulation may differ slightly from language to language; this rule will give us the long mid vowels from underlying ai and au. It will be a major rule and all those sequences which meet its SD will undergo it unless specifically marked for not undergoing it. The second rule will be needed for getting the diphthongs



ay and av; i.e., this rule will make the high vowel non-syllabic in the relevant context. This will be a minor rule and, thus, only those forms will undergo it which are positively marked for it. Again, although a particular language might have a slightly different version of this rule, broadly speaking such a rule will be required in all those languages which have the diphthongs ay and av. It follows from this that if the forms with surface diphthongs lose the positive marking for the rule, the result will be the mid vowels, as the MVR, or some variant of it, is a major rule and will apply to all forms meeting its SD unless otherwise stopped from doing so. And this is precisely what has happened in all these cases.

A question might be asked here regarding this recurrent supply of new diphthongs at various stages in these languages. That is, after the change has taken place, there will be no diphthongs left in the language. How does the language acquire new sets of these diphthongs? Then, why do these newly acquired diphthongs also not undergo monophthongisation? Regarding the first question, we can say that new sources were always available in these languages. For example, in late OIA the source was the early Vedic vridhhi diphthongs āi and āu, which became ai and au. In MIA, it is true that initially we do not find diphthongs as all the instances of Sanskrit diphthongs had become monophthongs.

However, in later MIA and early NIA, because of large scale borrowing from Sanskrit, these diphthongs again came back in the language. Then, in both MIA and early NIA, Sanskrit intervocalic aya and ava became ai and au respectively. For example, Skt nayana > NIA nain 'eye', Skt bhramara > NIA bhāvara > Hindi bhāurā 'black bee', early NIA kavana > Hindi kaun 'who', Skt avatāra > NIA autār 'incarnation', etc. In fact, while discussing glides and diphthongs (3.3), we noted the fact that in Hindi there is no difference in pronunciation between ai/au and aya/ava, excepting of course in the formal style. Thus, in both MIA and NIA there were always two ready sources for a fresh supply of diphthongs: Sanskrit borrowings with diphthongs, and Sanskrit intervocalic -aya- and -ava-. Also, in NIA at least, both the forms have existed side by side.

Regarding the second question, i.e., why did these new diphthongs also not become monophthongs, it is a well-attested fact that a particular sound change remains active in a language only for some time, after which it starts losing its intensity and gradually disappears completely. In this case we can say that the monophthong rule was active in the initial stages, i.e., it was transparent; later on, as new words with diphthongs started coming in, the rule started becoming opaque; i.e., a number of surface forms which were exceptions to this rule came into the

language. These exceptions were dealt with as exceptions, i.e., by marking them in the lexicon as not undergoing the monophthongisation rule. (For transparency and opacity of rules see Kiparsky 1971. According to Wang 1969 competing sound changes can also be a cause of such exceptions to a general rule.) Thus, although the rule remained in the grammar, it became restricted in its application. Consequently, the repetition of this change at intervals can be explained in terms of the rule's becoming more general or transparent once again.

Another question which might be raised here is that if ai and au become ē and ō respectively, then why is the process not reversible, i.e., why do we not get ai and au from ē and ō respectively anywhere? The answer is very simple. When the diphthongs become monophthongs, it is simplification of phonetic structure, as in place of two segments, we get only one. But if the monophthongs become diphthongs, it will be a complication as in place of one segment we shall get two. Furthermore, the historical development of IA is full of instances where a reduction in the number of surface segments has been effected precisely by the method of coalescence of two or more segments into one. Compare, for example, Skt sāṇdhya : Hindi sājh, 'evening', Skt satya : Hindi sach 'truth'. Skt matsya :

language. These exceptions were dealt with as exceptions, i.e., by marking them in the lexicon as not undergoing the monophthongisation rule. (For transparency and opacity of rules see Kiparsky 1971. According to Wang 1969 competing sound changes can also be a cause of such exceptions to a general rule.) Thus, although the rule remained in the grammar, it became restricted in its application. Consequently, the repetition of this change at intervals can be explained in terms of the rule's becoming more general or transparent once again.

Another question which might be raised here is that if ai and au become ē and ō respectively, then why is the process not reversible, i.e., why do we not get ai and au from ē and ō respectively anywhere? The answer is very simple. When the diphthongs become monophthongs, it is simplification of phonetic structure, as in place of two segments, we get only one. But if the monophthongs become diphthongs, it will be a complication as in place of one segment we shall get two. Furthermore, the historical development of IA is full of instances where a reduction in the number of surface segments has been effected precisely by the method of coalescence of two or more segments into one. Compare, for example, Skt sāṇdhya : Hindi sājh, 'evening', Skt satya : Hindi sach 'truth'. Skt matsya :

Bengali māchh 'fish', Skt. gardabha : Hindi gadhā 'donkey', Skt sthira : Hindi thir 'steady', etc. (The last two illustrate the deletion of segments as well.) Of course, it is not the case that the complication of surface structure is never found; epenthesis is the most common method of introducing a segment in the surface structure. But epenthesis is always used either to break a consonant cluster or to introduce an epenthetic glide between two vowels; the motivation in both the cases is to yield a less marked segment sequence. But this is precisely what monophthongisation does.

In summary, we can say that ai and au have become respectively ē and ō in MIA and in some NIA languages in quite a natural way, because the monophthongisation rule was always present in their grammars, and so, in every case, the change has meant the optimization of an already existing rule.

#### 5.2.6.4 Problems and questions in the synchronic evidence

While we are considering the dialect situation in NIA, we would like to examine two other languages/dialects, viz., WH and Avadhi. We shall try to show that the diphthongal analysis of the mid vowels being proposed here can explain certain facts about both of these languages in a very natural

Bengali māchh 'fish', Skt. gardabha : Hindi gadha 'donkey', Skt sthira : Hindi thir 'steady', etc. (The last two illustrate the deletion of segments as well.) Of course, it is not the case that the complication of surface structure is never found; epenthesis is the most common method of introducing a segment in the surface structure. But epenthesis is always used either to break a consonant cluster or to introduce an epenthetic glide between two vowels; the motivation in both the cases is to yield a less marked segment sequence. But this is precisely what monophthongisation does.

In summary, we can say that ai and au have become respectively ē and ō in MIA and in some NIA languages in quite a natural way, because the monophthongisation rule was always present in their grammars, and so, in every case, the change has meant the optimization of an already existing rule.

#### 5.2.6.4 Problems and questions in the synchronic evidence

While we are considering the dialect situation in NIA, we would like to examine two other languages/dialects, viz., WH and Avadhi. We shall try to show that the diphthongal analysis of the mid vowels being proposed here can explain certain facts about both of these languages in a very natural

way. This is not possible by either the previous analysis discussed here or by an analysis positing mid vowels in the underlying representations. We feel that the diphthongal analysis gains much in explanatory power as the facts about these two languages seem to be entirely unrelated with each other. For example, in WH we get æ and ɔ in place of ai and au of Hindi; in Avadhi we get yā and vā in place of ē and ō of Hindi. We shall take these up in the order stated.

As far as the mid vowels are concerned, there is no difference between WH and Hindi; both of them show the same mid vowels in the same forms. The difference is to be found only in the case of diphthongs: in place of Hindi diphthongs, WH shows low vowels, i.e., æ for ai and ɔ for au. (A number of linguists, e.g., Dixit 1963, Kelkar 1968, have taken WH as synonymous with Hindi; we, however, have preferred to take WH as synonymous with kharī Bōlī, a western NIA dialect on which both Hindi and Urdu are based. We might add that the diphthongal pronunciation is a characteristic of non-western Hindi as well as of non-western Urdu.) This difference between the two can again be very naturally and easily explained in terms of a difference between the rule components of the two languages rather than in terms of restructuring. In both WH and Hindi we have ai and au underlyingly for ē and ō respectively,

and MVR remains in the phonologies of both. In both it is a major rule. Hindi has a minor rule yielding the diphthongs. We can state this rule as follows.

PR16: Diphthong Rule (DIPH)

$$\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{hi} \end{bmatrix} \rightarrow [-\text{voc}] \quad / \quad \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ +\text{bk} \end{bmatrix} - \left\{ \begin{bmatrix} +\text{cons} \\ \# \end{bmatrix} \right\}$$

This rule turns a high vowel after a into a glide when it is immediately followed by either a consonant or a word boundary. As it is a minor rule, only those forms will undergo it which will be marked positively for undergoing it and negatively for MVR. (By convention, all forms are marked + for major rules and - for minor rules; thus exceptions to a major rule have to be specifically marked - for that particular rule and + for the minor rule which they undergo.) Now what has happened in WH is that the place of DIPH has been taken by another rule which coalesces ai and au into æ and ɔ respectively. We can state this rule as follows.

TR5: Low Vowel Rule (LVR (WH))

$$\begin{matrix} \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ +\text{bk} \\ -\text{long} \end{bmatrix} & \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{hi} \\ -\text{long} \\ \text{abk} \\ \text{arnd} \end{bmatrix} & \rightarrow & \begin{bmatrix} 1+2 \\ +\text{lo} \\ \text{abk} \\ \text{arnd} \end{bmatrix} \\ 1 & 2 & & \end{matrix}$$



and MVR remains in the phonologies of both. In both it is a major rule. Hindi has a minor rule yielding the diphthongs. We can state this rule as follows.

PR16: Diphthong Rule (DIPH)

$$\left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{hi} \end{array} \right] \rightarrow [-\text{voc}] / \left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ +\text{bk} \end{array} \right] - \left\{ \begin{array}{l} [+ \text{cons}] \\ \# \end{array} \right\}$$

This rule turns a high vowel after a into a glide when it is immediately followed by either a consonant or a word boundary. As it is a minor rule, only those forms will undergo it which will be marked positively for undergoing it and negatively for MVR. (By convention, all forms are marked + for major rules and - for minor rules; thus exceptions to a major rule have to be specifically marked - for that particular rule and + for the minor rule which they undergo.) Now what has happened in WH is that the place of DIPH has been taken by another rule which coalesces ai and au into æ and ɔ respectively. We can state this rule as follows.

TR5: Low Vowel Rule (LVR (WH))

$$\begin{array}{ccc} \left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ +\text{bk} \\ -\text{long} \end{array} \right] & \left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{hi} \\ -\text{long} \\ \text{abk} \\ \text{arnd} \end{array} \right] & \rightarrow \left[ \begin{array}{l} 1+2 \\ +\text{lo} \\ \text{abk} \\ \text{arnd} \end{array} \right] \\ 1 & 2 & \end{array}$$

But the lexical marking does not change; that is, LVR (WH) applies precisely to those forms which are marked [-MVR] and [+LVR(WH)]; these are the forms which in Hindi also are marked [-MVR] but [+DIPH]. The result will be naturally æ and ɔ in WH but ay and av in Hindi. For example, the underlying form /khait/ will have no marking in either dialect; thus in both MVR will apply giving us khēt 'farm'. But underlying /bail/ will be marked [<sup>-MVR</sup><sub>+DIPH</sub>] in Hindi, giving us bayl 'ox'; in WH, on the other hand, it will be marked [<sup>-MVR</sup><sub>+LVR(WH)</sub>], giving us bæl. (Details irrelevant to the present discussion have been omitted from the underlying representation.) Thus it is clear that the number of rules, in the present context, has not changed; in the place of one rule simply another has been substituted. That is, it is a case of both rule loss and rule addition.

It is possible to have a general rule for Hindi which says: if [-MVR], then [+DIPH]; similarly for WH: if [-MVR], then [+LVR(WH)]. This will mean that each relevant form will not have to be marked specifically [-MVR] as well as + for the relevant minor rule. As there is obviously an either/or condition on MVR and the relevant minor rule, there is clearly a generalization to be captured here; the proposed general rule does exactly this.

When we come to Avadhi as spoken in Unnao and a few other districts of Uttar Pradesh, we find that as far as the diph-

thongs ai, au are concerned there is no difference between Hindi and Avadhi, except for the fact that the rule lowering the glides (PR7, 3.3) is missing in Avadhi. For example, Compare H. [ba<sup>e</sup>l] : Av. [ba<sup>i</sup>l] 'ox', H. [a<sup>e</sup>sā] : Av. [a<sup>i</sup>s] 'like this', H. [a<sup>o</sup>r] : Av. [a<sup>u</sup>r] 'and', H. [kã<sup>o</sup>n] : Av. [kã<sup>u</sup>n] 'who', H. [pa<sup>e</sup>sā] : Av. [pa<sup>i</sup>sā] 'money', etc.

But when we come to the mid vowels we find that the Hindi mid vowels ē and ō correspond to yā and vā respectively in this dialect of Avadhi. Compare, for example, H. khēt : Av. khyāt 'farm', H. pēt : Av. pyāt 'stomach (n)', H. ghōRā : Av. ghvāRā 'horse', H. kōṭ : Av. kvāt 'coat', H. ēk : Av. yāk 'one', H. bēch : Av. byāch 'sell', H. khēl : Av. khyāl 'play (n,v)', H. ōr : Av. vār 'towards', H. rōk : Av. rvāk 'stop (v,tr)', etc. However, now examine these Hindi and Avadhi forms when a derivational suffix has been added to them; note that all the forms that we have given above are stems or base forms, both in the case of nouns as well as in the case of verbs. Consider, for example, H. and Av. khetihar 'farmer', H. and Av. bik 'be sold', H. and Av. ruk 'stop (v,intr)', H. ikkā : Av. ekkā 'ace', H. ghuRsavār : Av. ghoRsavār 'horse rider', H. khilārī : Av. khelārī 'player', Av. petavā 'stomach (dim.)', Av. koṭavā 'coat (dim.)', etc. (There is a slight simplification of the Avadhi forms as given here. For example, the final short u of the masculine forms has been left out. Then, Avadhi

normally shows r in place of Hindi R. Neither of these is relevant here.) It is clear that when we come to the derived forms, there is no difference between Hindi and Avadhi, whether the word has been derived from a noun or a verb. In other words yā and vā are found only in Avadhi base forms; in derived forms, i.e., after a derivational suffix has been added to the base, we get the short mid vowel; the short high vowels found in place of short mid ones in Hindi are due to the Mid Vowel Raising rule (PR10', 5.2.5).

To account for these alternations in Avadhi, if we represent the vowels in the base forms as underlyingly short high vowels, we need a rule to insert an ā after the high vowel for the base forms; after this GR (NR1, 3.1) can apply and turn the high vowel into a glide, thus giving us yā, vā. As GR is a natural rule, it will be as much a part of Avadhi phonology as of Hindi phonology. In case of derivation, we need a different rule to turn the high vowel into mid.

Positing mid vowels underlyingly does not improve matters. We still require the rule inserting a long ā after the mid vowels in the base forms; then we also require a second rule to turn the mid vowel into high. After this, GR can apply turning the high vowel into a glide. Alternatively

we can have a rule which will turn the mid vowel into a glide, this way we shall not need GR in this particular case. Either way we shall be able to get yā and vā in the base forms. In derivation, matters will be quite simple as SVS exists in Avadhi also; the application of SVS will get us the short mid vowels in the derived words.

We can also posit underlying ya and va, or more precisely, iā and uā. In this case we shall require an ā-deletion rule in place of an ā-insertion rule. This rule will, of course, be needed only in the course of word derivation. After this rule applies, we shall be left with the short high vowels; to get the short mid vowels from them, a rule like Flip will be needed as our revised Mid Vowel Raising will not be able to apply. As we have shown that Flip is not needed for Hindi, it will be a part of Avadhi phonology only.

As far as we can see, there are two questions involved here. First, since all the three proposed analyses account for the data equally well, which one do we choose? It is possible that more work on the language might make our choice easier; however, on the basis of the data given, the choice will be an arbitrary one. Second, is this peculiar fact about Avadhi related to Hindi in any way? As far as our knowledge of NIA languages goes, Avadhi is

unique in exhibiting this phenomenon. It is possible that this alternation belongs specifically to Avadhi, and as such we need not look for anything in common with Hindi. However, before accepting this we would like to see if we could find some explanation for this fact, which could relate Avadhi to other NIA languages, especially Hindi.

If we posit underlying āi and āu for yā and vā respectively in Avadhi, we can find precisely such an explanation. This language has added a rule of vowel metathesis which, operating on underlying āi and āu, gives us iā and uā respectively; the high vowel then is automatically turned into a glide by GR, and we can get yā and vā.. The vowel metathesis applies to base forms only, i.e., stems and roots. As far as the derived forms are concerned, there are two possibilities. First, we can get the SVS itself to delete the ā in this language; this would, of course, mean a reformulation of SVS for Avadhi. As this deletion will result in a lessening of the number of moras in the word, it can be quite properly subsumed under shortening. After this, as we shall be left with the short high vowel only, Flip or perhaps a High-Mid Rule will be needed for getting the short mid vowels. Alternatively, we can have MVR for Avadhi also, but constrain it to apply in word derivation only. After MVR has yielded the long mid vowels, SVS will

apply and shorten them. We might add that in the case of derived words in Avadhi, it is the short mid vowels which are generally found. Thus, this might be the better way.

However, the main point is that this analysis gives a reason for these 'idiosyncratic' alternations of this language; given underlying āi and āu, it is clear that yā and vā are possible, but kā and ṭā, for example, are not. As far as the diphthongs are concerned, we require another minor rule for shortening the long ā. Thus, the entire output is taken care of. Incidentally, this is yet another instance where a language differs from its related languages because of a change in the rule component, in this case by the addition of a rule of vowel metathesis. We can formally state the vowel metathesis rule as follows.

TR6: Vowel Metathesis (Rule for Avadhi)

$$\begin{array}{ccc}
 \left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{bk} \\ +\text{lo} \\ +\text{long} \end{array} \right] & \left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{hi} \end{array} \right] & \begin{array}{c} X \\ \left[ \begin{array}{l} -\text{deriv.} \\ \text{suffix} \end{array} \right] \end{array} \rightarrow \begin{array}{ccc} 2 & 1 & 3 \end{array} \\
 1 & 2 & 3
 \end{array}$$

The rule as formulated will apply only when no derivational suffix follows. One more important point to be noted in connection with this dialect and the ē/ō dialects discussed earlier is that in both cases the change has resulted in the optimization of an already existing rule: MVR in the case of

ē/ō dialects, GR(3.1) in this case.

We might mention another interesting fact regarding OIA phonology on the one hand, and the phonology of Hindi and other NIA languages discussed here, on the other. In OIA, āi and āu gave us the diphthongs; ai and au gave us the mid vowels, and there was a rule shortening the long ā in the case of diphthongs. In Hindi, there seems to be a slight restructuring as we have only ai and au for both the mid vowels and the diphthongs, and the rule shortening the ā has been lost. However, it seems that in dialects this rule still exists in some form; witness, for example, pronunciations like [mael], [laen] for the Hindi [māil] 'mile', [lāin] 'line'. But it is possible to posit āi and āu underlyingly for Hindi as well; in fact, as we do not have any specification for [long] for a in our MVR, it will turn both āi/āu and ai/au into ē/ō, and we do not know of any facts about the language which would militate against this. If this is the case, then in Hindi too āi and āu will yield the mid vowels and ai and au the diphthongs; the OIA rule shortening the ā has been lost in either case. From this point of view, no restructuring has taken place, only a change in the rule component.

### 5.2.7 Ordering the Rules

We have thus four important rules which operate on underlying vowel sequences; these are MVR, GR(3.1), Post-



Glide a-deletion (PR6, 3.3), and DIPH. Of these, MVR turns a sequence of a and a following high vowel into the corresponding long mid vowel; GR turns a high vowel into a glide when the high vowel is immediately followed by another vowel; PR6 deletes an a when it is immediately preceded by ay or av, and DIPH turns a high vowel into a glide when the high vowel is immediately preceded by a. There is no ordering problem between MVR and DIPH as there is an either/or condition imposed on them. But there can be ordering problems among the other rules, and so we would like to examine them a bit further; we would also like to examine whether there could be any change in the formulation of these rules. For the sake of convenience we reproduce the rules below.

TR4: MVR

$$\begin{array}{ccc}
 \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ +\text{bk} \end{bmatrix} & \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{hi} \\ \alpha\text{bk} \\ \alpha\text{rnd} \end{bmatrix} & \longrightarrow & \begin{bmatrix} 1+2 \\ -\text{hi} \\ -\text{lo} \\ \alpha\text{bk} \\ \alpha\text{rnd} \\ +\text{long} \end{bmatrix} \\
 1 & 2 & & 
 \end{array}$$

NR1: GR

$$\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{hi} \end{bmatrix} \longrightarrow [-\text{voc}] \quad / \quad \text{---} \quad \begin{bmatrix} +\text{voc} \\ -\text{cons} \end{bmatrix}$$

PR6: Post-Glide a Deletion

$$\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ +\text{bk} \\ -\text{long} \end{bmatrix} \rightarrow \emptyset / \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ +\text{bk} \\ -\text{long} \end{bmatrix} \begin{bmatrix} -\text{voc} \\ -\text{cons} \\ +\text{hi} \end{bmatrix} -$$

PR16: DIPH

$$\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{hi} \end{bmatrix} \rightarrow [-\text{voc}] / \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ +\text{bk} \end{bmatrix} - \left\{ \begin{bmatrix} +\text{cons} \\ \# \end{bmatrix} \right\}$$

The problem is this: Suppose we have an underlying sequence aia, then what will be the output? We can think of two possibilities. First, MVR applies giving us ēa; after this, none of the other three can apply. Thus, this order of rules is a bleeding order. (For feeding and bleeding orders see Kiparsky 1968a, and with a different terminology, Chafe 1968.) Second, GR applies first giving us aya. Now MVR cannot apply, though PR6 can, giving us ay. Either way we get a bleeding order, and out of MVR and GR it seems that only one can apply.

The matters, however, are not altogether bad. When we look at the data from the language, we find that a long mid vowel is never followed by a vowel, but always by either a consonant or a word boundary, precisely as is the case with DIPH. It follows from this that MVR as it has been formulated by us is too strong and has to be

constrained. It is clear that the rule never applies to an underlying, say, aia, but only to aiC/##; thus what is required is a change in the formulation of MVR. Consequently, we reformulate MVR as follows.

TR4' : MVR(R)

$$\begin{array}{ccc}
 \begin{array}{|l|} \hline +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ +\text{bk} \\ \hline \end{array} & \begin{array}{|l|} \hline +\text{voc} \\ -\text{cons} \\ +\text{hi} \\ \text{abk} \\ \text{arnd} \\ \hline \end{array} & \begin{array}{|l|} \hline [+ \text{cons}] \\ \# \\ \hline \end{array} \rightarrow \begin{array}{|l|} \hline 1+2 \\ -\text{hi} \\ -\text{lo} \\ \text{abk} \\ \text{arnd} \\ +\text{long} \\ \hline \end{array} \quad 3
 \end{array}$$

1                      2                      3

Diphthongs proper, whether in Sanskrit words or native words, are also never followed by a vowel. By diphthongs proper we mean those diphthongs which come from an underlying ai/au and not from an underlying aia/aua, after the application of GR and Post-Glide a-deletion. After this reformulation of MVR there will be no ordering problems. Given underlying aia/aua, only GR obligatorily and Post-Glide a-deletion optionally can apply, in this order. Given underlying ai/au, only MVR or DIPH can apply, as there is an either/or condition on the two. This division of these four rules into two sets of two each is not only neat but also accords well with the facts of the language. There is obviously a formal similarity between MVR and DIPH. Disregarding the glides in the diphthongs, GR is the only source for high glides in the language; consequently the pairing of GR and PR6 is also quite natural.

### 5.2.8 Is DIPH Necessary?

So far we have proceeded on the assumption that all of these four rules are necessary; in what follows we would like to examine this assumption. More precisely, we would like to examine whether DIPH is necessary.

While discussing the pronunciation of glides (3.3) we saw that a large number of Sanskrit words with aya/ava are reduced to ay/av respectively in Hindi, unless the words are pronounced as Sanskrit words. For example, nayana 'eye' is pronounced both as nayan and nayn, avasara 'chance' is pronounced both as avasara and avsara. (We disregard the deletion of the word final a here.) In fact, our Post-Glide a- deletion rule (PR6, 3.3) was formulated to take care of just such a situation; the rule by reducing aya/ava to ay/av has precisely the effect of giving a diphthongal pronunciation of the former. This, of course, is for those Sanskrit words which have original aya/ava only. But if we represent the proper diphthongs also as underlyingly aia/aua, we can do away with DIPH, as now GR and Post-Glide a- deletion, applying in this order, will always yield the correct output. By making the latter rule optional for Sanskrit words with original aia/aua, and obligatory in the rest of the cases, we can ensure that in the first case we shall get both aya/ava

and ay/av, but in the second case we shall get ay/av only; and this is precisely what is needed.

Although this new analysis is able to give us the correct output, and that too with a fewer number of rules, we do not think that it is correct. The use of marking in this case again seems to us like a gimmick for the 'optimization' of a particular rule and has no basis in the facts of the language. In fact, it is spurious optimization. Not only this. The use of Post Glide a- deletion rule to delete an a which never surfaces is equivalent to absolute neutralization. Although Kiparsky 1971 does not banish absolute neutralization from phonological theory, as did Kiparsky 1968b, he allows it only if the abstract segment is needed for more than one rule. As far as we know, this is not the case here. The fact that nayana 'eye' and avasara 'chance' can be pronounced as both nayan/nayn and avasara/avsar, whereas bayl 'ox' and kavn 'who' can never be pronounced as bayal and kavan is not because the Post-Glide a- deletion is optional in the former case and obligatory in the latter. It is because there is no a after the glide to be deleted in the latter. These two words are underlyingly baila and kauna, whereas the other two words are underlyingly naiana and auasara. It is true that this way we need a rule just to account for such words; it is also true that we

shall have two sources for the same phone. However, we do not think that either is particularly damaging to our analysis. While discussing vowel nasalization we took the same stand and specified some nasalised vowels as underlyingly [+nas], but derived the others through the Vowel Nasalization Rule (4.2.2.1). Here also we would like to do the same thing and derive surface diphthongs from two different underlying sources. We might add that DIPH will have a fairly large number of forms to apply to, as diphthongs proper are found in both Sanskrit and Hindi words. Thus, we keep DIPH as it is.

We think that we have been able to show that a diphthongal analysis of mid vowels in Hindi is superior to any other analysis that we have examined here. We also think that it has been possible to do so simply because we considered the question in the light of evidence from other dialects of Hindi as well as other NIA languages. It will also be clear that while doing so, we never ignored or underplayed the importance of synchronic data; on the contrary, our aim was to account for it. We also thought that if our analysis could account for facts in the related languages and dialects as well, it would be much better, an aim that we mentioned earlier as being a desirable one. Nor can our analysis be really called abstract. As far as we understand it, an abstract analysis is one which makes

use of absolute neutralization; there is no absolute neutralization involved in our analysis. In fact, this was one of the reasons why we preferred to have DIPH. As far as the motivation for having underlying diphthongs is concerned, we think that we have said enough about it in the preceding pages, and we feel that we have proved our case.

## CHAPTER 6

### ON WORD FINAL a

#### 6.1 BACKGROUND

In the preceding pages we have referred to the a-deletion rule a number of times but have not discussed it. The rule has been discussed by several writers, e.g., Kellogg 1955, Misra 1967, Srivastava 1969, Pray 1970, Narang and Becker 1971, Ohala 1972, 1974a, 1974b, Pandit 1976, Bhasin 1978, and we have more or less accepted their formulation of it. However, all of these have discussed this rule as a schwa-deletion rule applying to a word medial a. In this they have been guided by phonetic facts as normally a short a is not found in the word final position in Hindi. On the other hand, our claim is that all those Hindi words which end in a consonant phonetically have a short a in the word final position in the underlying representation; this a is deleted by a rule later on. It is this rule with which we shall be concerned in this chapter. Next, we will show that if the rule is formulated in terms of distinctive features, it becomes a very general rule applying in almost all NIA languages in some form or



the other. Although we will not be concerned with the deletion of a in the word medial position, we shall have occasion to refer to this context, hence we give the rule as formulated by Ohala (1974:10) below. We retain the numbering of Ohala.

$$(20) \text{ } \text{ə} \rightarrow \emptyset / \text{VC}_1^2 \text{ — CV} / \left\{ \begin{array}{l} \left[ \begin{array}{l} +\text{loan} \\ +\text{casual speech} \end{array} \right] \\ \left[ \begin{array}{l} +\text{normal} \\ \text{tempo} \end{array} \right] \end{array} \right\}$$

- Condition 1: There may be no morpheme boundary in the environment to the left.  
 Condition 2: The output of the rule may not violate the sequential constraints of Hindi.  
 Condition 3: The rule applies from right to left.

We might add that the second condition has also been mentioned by Pandit 1976 and Bhasin 1978.

## 6.2 POSITING WORD FINAL a

The justification for positing a word final a is not as simple as that for the medial a, for the simple reason that whereas the medial a surfaces in some of the related forms and is deleted in only some others, the word final a normally never surfaces in speech. Thus the deletion in this context becomes very much like absolute neutralization and as such, if not indefensible, is a very costly feature of grammar indeed (Kiparsky 1971). Consequently, the evidence justifying the positing of a final

a will be presented in the following manner. First, we shall show that a final a enables us to explain certain facts about Hindi phonology in a natural way, which is not possible in the absence of a final a. Second, we shall show that there exist certain contexts in which the final a does surface, and, we might add, there are quite a few of these contexts. We shall discuss these in the order given.

### 6.2.1 Facts of Hindi Phonology

#### 6.2.1.1 Flaps in Hindi

The presence of voiced retroflex flaps, both unaspirated and aspirated, is a characteristic of several NIA languages, especially the modern literary ones, and Hindi is no exception. In fact in Hindi even the retroflex nasal, found in Sanskrit words, has become a nasalised retroflex flap. We might mention that Ohala has shown on the basis of experimental evidence that, in Hindi, nasalisation can pass through R, y, v and h (1975:323). As far as the historical development of these sounds is concerned, there is no controversy; the intervocalic voiced retroflex stops of Sanskrit have become flaps in Hindi, Bengali and several other NIA languages. The problem arises only when we come to the synchronic grammar.

The problem is simply this. Although there are many words in which even synchronically the flaps are intervocalic,

e.g., ghōRā 'horse', kaRī 'a link', gāRī 'carriage', būRhā 'old (m,sg)' etc., there are a number of words for which this observation does not hold. First, we have words like laRkā 'boy', gaRh 'fort', saRnā 'to become rotten', paRhna 'to read' etc., in which a flap, although present, is not intervocalic. Second, we have words like niḍar 'fearless', kuḍaul 'crooked', bēḍaul 'out of shape' etc. in which although the voiced retroflex stop is intervocalic, it is not flapped. Because of evidence like this traditional phonemic analyses like Dixit 1963 and Kelkar 1968 have treated flaps and voiced retroflex stops as separate phonemes. Dixit 1963 gives marginal minimal pairs like Skt. uḍū 'star' and Hindi uRū 'may I fly', English Rhoda and Hindi rōRā 'stone' in support of the separate phonemic status of the two. Kelkar (1968:40) examines the question in greater detail, especially the argument that as the voiced retroflex stop is found only word initially, whereas the flap is found either intervocalically or word finally, the two are in complementary distribution, as such they can be subsumed under one phoneme. He rejects the argument on the basis of words like niḍar 'fearless' where the stop is intervocalic.

Although it is true that in a word like niḍar the stop is intervocalic and in a word like ghōRā the flap is intervocalic, it is equally true that for any native speaker

without a formative boundary before it. English words like rhoda, Ada etc. and Sanskrit words like udū 'star' etc. can be marked as exceptions to the rule. We might mention that the frequency of a word like udū is very small and all those Sanskrit words which have the slightest frequency in Hindi and have intervocalic stops like, e.g., tadit 'lightening', pīḍā 'pain', krīḍā 'sport (n)' etc., regularly undergo flapping and are pronounced as taRit, pīRā, krīRā respectively. We might formulate the rule as follows.

PR17: Flapping Rule

$$\left[ \begin{array}{l} +\text{obs} \\ +\text{ab rel} \\ +\text{voi} \\ +\text{ret} \end{array} \right] \longrightarrow [+flap] / \begin{array}{c} V \quad X - V \\ \left[ \begin{array}{l} -\text{seg} \\ -\text{WB} \\ -\text{FB} \end{array} \right] \end{array}$$

However, this rule will not account for flaps in words like gaRh 'fort', mōR 'turning', rēvaR 'flock', lāR 'affection' etc., in which the flap is word final and as such not intervocalic. One way to account for such words would be to remove the second vowel from the SD of our rule, that is, make the rule apply post-vocalically rather than intervocalically. In this way, no doubt, we shall be able to take care of the words given above, but we shall not be able to account for words like gaḍḍhā 'pit', aḍḍā 'a place', etc., where the retroflex stop is undoubtedly post-vocalic

but still is not flapped. We might try to remedy the situation by making the rule apply when the stop is either intervocalic, as before, or between a vowel and a word boundary, i.e., the second context is  $VX - \#$ . Undoubtedly this revised rule with its two contexts will be able to account for all of the cases given so far; however, the second context of the rule seems somewhat strange to us, to say the least. The intervocalic position seems a natural position for flapping; we can take the flapping of intervocalic t in American English as an example. But by having the contexts  $VX - V$  and  $VX - \#$  which can be reduced to  $VX - \left\{ \begin{matrix} V \\ \# \end{matrix} \right\}$ , the rule is making the claim that there is something in common between a vowel and a word boundary. Now although Lass 1971 has claimed that boundaries have some features in common with obstruents, nobody has suggested that boundaries have anything in common with vowels. Even if we reject this implication, we are still saddled with two contexts for the same process which are in fact absolutely independent of each other; as far as our knowledge goes, such contexts are not very frequent in natural languages.

The best solution of course would be if we could find some vowel that occurs finally in these words in the underlying representation but is later deleted to give us

phonetically C-ending words; in fact this could be done for all those words which phonetically end in a consonant, whether this consonant is a flap or any other. And such a vowel exists in a. We suggest that all those words which end in a consonant phonetically end in a short a in the underlying representation, and this a is later on deleted by the a-deletion rule. This immediately makes our Flapping Rule simpler as well as more natural: simpler in the sense that instead of having two contexts we have only one context left, the intervocalic context, and more natural in the sense that the intervocalic context is naturally conducive to flapping. As we are positing this a for all C-ending words and not just for those which have a flap in the final position, we cannot be 'accused' of resorting to this way just to make our rule more elegant. In any case there are a number of other arguments as well in support of this final a; however, before we take up those, we would like to discuss one more class of words in the context of flaps.

This class consists of words like mūḍēr 'cornice', ūdēlnā 'to pour' etc., in which the pre-stop vowel is nasalised, but there is no flapping. Basing ourselves on words like these alone, we might say that the rule does not apply if the vowel to the left of the stop is [+nas]. However, an alternative pronunciation of these and similar

words with flapping exists, i.e., mūRēr and ūRēlnā respectively. In view of this, it seems to us that for some speakers the rule does not apply if the vowel before the stop is nasalised; for other speakers there is no such constraint. There is another explanation possible as well. According to Kelkar (1968:33) and Ohala (1975:329,n3,4) when a nasalised vowel is followed by a voiced stop, there is always a short homorganic nasal present between the two giving us a pre-nasalised stop. If so, the vowel is followed by two and not one consonant; thus flapping is automatically blocked as the stop ceases to be intervocalic.

#### 6.2.1.2 Script and pronunciation

Let us now pass on to some other arguments in favor of the word final a. First, as far as writing is concerned, all such words are written with the short final a. Attributing this to the syllabic nature of the script in which every consonant letter stands for the consonant and the short a is not very conclusive, as the script has a diacritic mark for signifying a consonant alone. In fact, this mark, if needed, is regularly used in writing Sanskrit, which uses the same script as Hindi. Second, it is not the case that the final a is never pronounced in Hindi. Although in normal speech the words end in a consonant, the a does surface in slow speech or in a spelling pronunciation. It

also surfaces in songs whenever it is required by the metre or the rhythm. Any random survey of Hindi film songs will prove this beyond doubt. Some evidence from Punjabi regarding this question might not be out of place here. Punjabi and Hindi share a number of proper names, both for men and women. The words relevant to our purpose are those which end phonetically in a consonant in Hindi. But in Punjabi when these names are used vocatively, i.e., in the vocative case, the final a is always pronounced. Thus, the Punjabi vocatives [bhupindara], [satindara], [sarōṛja] etc. regularly correspond to Hindi [bhupindar], [satindar], [sarōṛ] respectively. In writing, of course, they are written with the final a in both the languages. In what follows, we consider some more examples where the positing of a word final a simplifies matters.

If we choose to go by the phonetic representation alone, we shall have to represent words like vichāra 'thought', phala 'fruit', darśana 'seeing (n)', dhana 'money', para 'the other', bhāva 'emotion', samaya 'time', etc. without the final a, i.e., as vichār, phal, darśan, dhan, par, bhāv and samay respectively, as they are normally written in Roman transcription of these words in Hindi. Note that all these are Sanskrit words but quite common in Hindi. There also exist a number of other fairly common



Sanskrit words like arpan 'offering (n)', arjan 'earning (n)', adhīn 'dependent', abhilāṣī 'desirous', anusār 'according to', atirēk 'excess' etc., which begin with a short a. (We have left out the final a in the case of C-ending words in the list.) These words can be compounded with the former giving us compounds like vichārādhīn 'under consideration', darśanābhilāṣī 'desirous of seeing', dhanārjan 'earning money', parādhīn 'dependent on others', bhāvātirēk 'excess of emotion', samayānusār 'according to time', etc. Now the lengthening of the final a of the first word in all these and similar cases is due to the sandhi rule of Sanskrit according to which two similar vowels coalesce to give the corresponding long vowel. This way matters are quite simple. But if we take pronunciations like [vičār], [dhan], [samay] etc. as representing the underlying representation of these words, then the compounds given above can only be explained on the following basis. Given word sequences like vichār+adhīn, dhan+arjan, samay+anusār etc., there is a rule in Hindi which lengthens the initial a of the second word. This rule applies only when the word is the second member in a compound of which the first word ends in a consonant. Note that we cannot say that the initial vowel is lengthened, as in dhan+upārjan, for example, the result is not dhanūpārjan but dhanōpārjan 'earning money'. And there are not just a few such cases.

Then, what could be the motivation for a rule lengthening the a, and changing u to ō? We do not know of any language in which such rules are found. In fact words like vichāradhīn, samayanūsār, dhanarjan, dhanupārjan, etc. are perfectly pronounceable and break no constraints, sequential or otherwise, of Hindi. Consequently, a large part of Hindi phonology will have a number of similar absolutely unmotivated rules. Dismissing these words as Sanskrit and non-native will not do. The percentage of Sanskrit words in Indian languages, either IA or Dravidian, is quite considerable. The fact that the complicated stress rules for English as given by Chomsky and Halle (1968) are based on the romance vocabulary does not make them non-English rules.

The final a also surfaces if it is preceded by a high vowel and a glide, e.g., in priya 'dear', rājsūya 'a sacrifice', nīva 'foundation', kṣatriya 'the warrior caste', etc. It is to be noted that the first word is never pronounced without the final a, although the glide may be dropped; even the tadbhava form pia has the final a.

The final a also surfaces if it is preceded by two consonants, e.g., in putra 'son', ratna 'jewel', bhakta 'devotee', anna 'foodgrains', vākya 'sentence' uparōkta 'above mentioned', daitya 'demon', umra 'age', sabra

'patience', etc., the final a is never deleted. We may mention that the last two are Arabo-Persian words; the others are Sanskrit.

So far we have concerned ourselves with mainly Sanskrit words in support of our contention that a final a is present in all those words at the underlying level which phonetically end in a consonant. Although the words that we selected are fairly common, it cannot be denied that they are Sanskrit tatsamas. We would now like to present some examples from native Hindi words where the final a always surfaces in one form or the other. The majority of Hindi verbs are of the type kar-nā 'to do', phisal-nā 'to slide', ghūm-nā 'to wander', etc., that is, phonetically the verb stem or root ends in a consonant. As opposed to these, there are verbs like kahe-nā 'to say', rahe-nā 'to live', sahe-nā 'to endure', bahe-nā 'to flow', where the verbal root ends in a short e, which is preceded by h. This short e preceded by h is derived from an underlying a by a rule (PRL, Section 2.2.1) which is absent in EH; in EH, in the place of this e we get a, e.g., kaha-nā 'to say'. On the basis of this we have to posit an underlying final a for these verbs at least, and the same will have to be done for words like griha 'house', grahe 'planet', sahe 'together', tahe 'depth', šahe 'checkmate'

dahe 'tank', sulahe 'armistice', subahe 'morning', tarahe 'kind', and some others. (Phonetically aha can be realised as [a<sup>e</sup>h] or [e<sup>h</sup>].) No doubt this group consists of a small number of words; yet the generalisation regarding this group seems to be that in Hindi if the two segments preceding the last one in a word happen to be ah, then the last segment can only be a. We might add that if these two segments are consonants then also the last segment can be only a. This will have to be modified further. As other vowels, e.g., ā, ī, ū etc., also can be the last segment after either h or CC, we have to say that if in the contexts hV# or CCV#, the V is non-high, non-mid, non-front, and short, then it must be a! But this amounts to saying nothing. In fact, it would be much better to insert an a by a rule. But even this will not end our problems. For example, if a vowel is inserted, it is epenthesis. Now epenthesis is always used to simplify a consonant cluster, but there is no consonant cluster to be simplified in the context h-#; if C# is possible, then why not h# or ah#? That is, our insertion rule will be totally unmotivated; note that there is nothing in common between the two contexts CC# and ah#. We might try to save the situation by making the rule more general. For example, why not get all the a's by this rule? That is, if the feature

matrix of a vowel segment is not fully specified, then the rule will change it into that of a. This way the rule will have many more forms to apply to; unfortunately, the rule remains as unmotivated as before; in fact it becomes worse. Stanley 1967 has criticised this kind of segment insertion as spurious simplification.

To summarise. Whether we posit a final a for all phonetically C-ending words or not, we have to posit a final a in the following cases:

1. All those words which end in a high vowel followed by a glide, e.g., priya 'dear', etc.
2. All those words which end in a consonant cluster, e.g., putra 'son', bhakta 'devotee', etc.
3. All those words which end in ah, e.g., kahe 'say', tahe 'depth', etc. (The final e comes from a; see 2.2.1.)

If apart from these three cases, we do not posit a final a in the case of words phonetically ending in a consonant, we have to say something like the following. In Hindi, a word final a occurs underlyingly only in the three cases mentioned above. If apart from these cases, a is found word finally, then it is due to:

1. Sanskrit influence.
2. The influence of the script.
3. The speaker's idiolect.

4. Requirements of the metre or rhythm.

5. Individual cussedness.

(The list is not exhaustive. We disregard the complication of the rule component as being of no relevance here.)

We can recall a number of strange statements in linguistics, especially when rival theorists have been debating in print, but this is by far the strangest.

On the other hand, if we posit a word final a for all those words which phonetically end in a consonant, no difficulties need arise. We do not even need a new rule; we simply have to extend the scope of an already existing rule ((20), 6.2) by adding an extra context to it. Of course, we shall have to restrict the rule from applying to an a if ah immediately precedes. But this restriction will have to be there whether we posit a final a or not, because a is never deleted if it is preceded by ah. Consider, for example, alternations like chamak 'brilliance' / chamakā 'shone (m,sg)', kamal 'lotus' / kamalō 'lotuses (obl,pl)', etc., where the a, underlined in the first alternant, is deleted in the second, as a syllable has been added. In precisely similar circumstances, a is not deleted if it is preceded by ah, e.g., maheak 'fragrance' / maheakā 'became fragrant (m,sg)', mahel 'palace' / mahelō 'palaces (obl pl)', where e, which comes from a, is not deleted in the second

alternant. In fact, as far as Hindi is concerned, no restriction is required to be placed on the rule; we simply have to let the rule changing a to e (PR1, 2.2.1) apply before the a-deletion rule. Thus, when the rule applies the a will have been already changed to e. But EH will require this constraint on the a-deletion rule, as EH does not have PR1.

We have thus given a number of arguments and examples in favor of positing a word final a for the phonetically C-ending words in Hindi. Although none of these taken alone might be conclusive proof in support of our position, we think that the combined weight of all of them is sufficient to justify our stand.

### 6.3 GENERALISATION OF THE RULE

We said in the beginning of this chapter that the rule deleting a short a has been discussed by several writers as the a-deletion rule (see references there); as such a question regarding the nomenclature can arise. That is, are there any advantages in calling the rule the a-deletion rule, as we have done, or are the two simply notational variants of each other? If the latter, then our decision to call the rule the a-deletion rule becomes a matter of personal preference and as such cannot have any

theoretical significance. In what follows we shall attempt to show that this is not the case and that our formulation of the rule, to be given below, enables us to achieve a much greater degree of generalization than the other one. One difference that can be pointed out in this regard is that whereas a refers to the physical phonetic level, a refers to the systematic phonemic level in the sense of Chomsky 1964a. This does not apply to Srivastava 1969, who has taken a as equivalent to a and a as equivalent to a; thus in his formulation of the rule he has given  $\begin{bmatrix} -\text{tns} \\ +\text{lo} \end{bmatrix}^{\text{v}}$  as the feature matrix for a (Srivastava 1969: 919). But it is not explicit in the other cases whether a has been taken as an underlying segment or as a derived one, although on the basis of the data given by, for example, Ohala 1974b, it seems that it is derived from an underlying short a. In our opinion also a is derived from an underlying a which is unstressed at the surface level. This is not at all surprising as a is the reduced vowel par excellence; in English, for example, every vowel can be reduced to a, provided it loses its stress to a certain extent. It also follows from this that a vowel reduction rule reducing all or some unstressed vowels to a will be a part of the phonologies of many languages.

The rule under discussion is a fairly old one. Misra 1967 has shown its existence in Old Hindi; McGregor 1968



has shown that in the 18th century it existed in Braj, a western NIA language quite close to Hindi. Pandit 1976 has shown its existence in Old Gujarati as well as Modern Gujarati; he has also shown that the word final context for a-deletion is historically prior to the word medial context. Kellogg 1955 (first edition 1875), Guru 1963 (first edition 1921), Sharma 1975, and other Hindi grammarians following the traditional or not-so-traditional framework have also mentioned the rule. Kellogg has mentioned a very interesting fact about a-deletion in the medial position. According to him, in the medial position there is always a schwa-like sound instead of the normal a, though in the final context we get no vowel at all (Kellogg 1955:11). This shows that even by the 1870's the a-deletion in the medial context was not firmly established. Perhaps this is the reason why there seem to be so many exceptions to any formulation of this rule. (See Ohala 1974b for a discussion of these.) The rule is also found in other NIA languages like Marathi, Punjabi etc.

The motivation for this rule in NIA is not difficult to find; it is a part of the general simplification process witnessed in NIA in its development from OIA, of which the reduction in the number of syllables of a given word is one aspect. Note that the rule in effect deletes one syllable. More concretely, the motivation can be found in

the preferred syllable structure of Hindi and other NIA languages. The preferred syllable structure in Hindi is either  $C\bar{V}$  or CVC, CV being found mostly word initially. (See Chapter 7 for some discussion of the subject.) Note, for example, the lengthening of short high vowels word finally (4.1.1.2), and the operation of CVL (4.1.1.1) both of which result in a  $C\bar{V}$  structure; the a-deletion rule, on the other hand, yields a CVC structure. Pandit 1976 has given stress shift in NIA as a motivation for the rule as well.

As the motivation of the rule is due to an NIA phenomenon, it is not surprising that it is found in all those NIA languages which have a, i.e., all the non-Eastern NIA languages. In the Eastern NIA languages a is normally not found as it has become  $\text{ɔ/o}$ ; obviously the rule as such is inapplicable there. On the other hand, the reduction in the number of syllables being an NIA phenomenon, it is possible that the Eastern NIA languages have some other rule which performs precisely the same function as does the a-deletion rule. It also follows that because of the correspondence between a in the non-Eastern NIA languages and  $\text{ɔ/o}$  in the Eastern NIA languages, if such a rule exists, it will in all probability take the form of an  $\text{ɔ/o}$  deletion rule. And when we look at data from, for example, Bengali, we find that this is precisely

the case; in other words wherever a is deleted in Hindi, ɔ is deleted in Bengali, and wherever a is not deleted in Hindi, because of the inapplicability of the a-deletion rule, ɔ is also not deleted in Bengali. (Henceforth we shall be using ɔ as a cover dymbol for both ɔ as well as o.) For example, word finally if we do not get a in Hindi, we do not get ɔ in Bengali, e.g., H. kamal : Beng. kɔmɔl 'lotus', H. phal : Beng. phɔl 'fruit'; if we get the word final a in Hindi as there is a CC cluster preceding it, in Bengali we get ɔ, e.g., H. basanta : Beng. bɔʃɔntɔ 'spring', H. sudīpta : Beng. ʃudīptɔ 'well-illuminated', and so on. (The change of Hindi s to Bengali ʃ is irrelevant here.)

The examples given above, and a number of others that we have not given for want of space, show clearly that it is necessary to have an ɔ-deletion rule for Bengali, although the frequency of deletion in Bengali is somewhat less. In Oriya, another non-Eastern NIA language, the rule has only marginal application in the standard dialect. Interestingly enough, the deletion there seems to be syntactically governed, as it applies only to nouns in the vocative case and verbs in the non-honorific imperative form. However, in the western dialect of Oriya, the rule applies much more freely; note that this dialect is spoken in areas bordering Madhya Pradesh, a Hindi-speaking state. Traditional Oriya grammars recognise this fact as well as

the fact that a-deletion is a characteristic of Hindi. As the contexts for the applicability as well as non-applicability of the ɔ-deletion rule will be the same as those of the a-deletion rule, it would be much better if the two could be expressed as one rule. Because then we shall get a very general rule which apart from being a part of the grammar of these specific languages will also be a part of the grammar of NIA. And the moment we use distinctive features for representing a and ɔ, our motivation becomes clear. According to the SPE, a and ɔ are minimally different in the sense that where a is [-rnd], ɔ is [+rnd], all other specifications being the same. In other words if we do not mention the feature [rnd], the vowel can be either a or ɔ; as these two vowels are never found together in any NIA language, we can express both the a-deletion as well as the ɔ-deletion rules as a short low back vowel deletion rule. It will also be clear now why we preferred to call our rule a-deletion rule rather than the ə-deletion rule, as it has been called by others. In terms of distinctive features a and ɔ are minimally different, but ə and ɔ are not; thus the greater degree of generalization that we have been able to achieve has been possible because first, we have used the underlying segments and second, because we have used the distinctive features rather than the segments themselves in our rule. It is one more proof, if it be needed, for the

assertion of generative phonology that it is the distinctive features which are important in phonological processes and that linguistically significant generalizations can be most naturally and effectively captured only through the use of distinctive features in rules. (On the controversy regarding the use of distinctive features in phonological rules, see Householder 1965 and the rejoinder by Chomsky and Halle 1965.)

We might add that even if it be proposed that Bengali and other ENIA languages do not have an underlying ɔ but rather a which is later tuned into ɔ by a rule, it does not make any difference to what we have said here; the rule remains a short low back vowel deletion rule. Also that the presence of ɔ as well as o does not make any difference either. The raising of ɔ to o can be easily taken care of by a rule, presumably a late level phonetic rule.

It might be said that by formulating the rule as an a-deletion rule, or a short low back vowel deletion rule, we are adding one more rule to our grammar, the rule which will change a to ɔ. But this is a rule which we are going to have in the grammar anyway. For example, words like katnā 'to be spun' / katānā 'to get spun', hatnā 'to get aside' / hatānā 'to cause to get aside', gayā 'went (m,sg)',

paRḥā 'studied (m,sg)', baḍal 'change (v)' etc. are realised phonetically as [kátnā]/[kaṭānā], [háṭnā]/[haṭānā], [gaṽā], [paR<sup>h</sup>ā], [baḍāl] respectively. It is clear that if a loses its stress, it is reduced to ə, a very common rule in many languages. On the other hand, by calling the rule a short low back vowel deletion rule, we are claiming that the rule deletes a vowel which is underlyingly a in Hindi, no matter what its phonetic realisation, be it [ə] or [ʌ] or [ɪ]. In some eastern dialects of Hindi, a is realised as ɪ; the proposed rule will cover these dialects also. The rule can be stated as follows. (We have taken the word final context only as this was the context we were concerned with here.):

PR18: Short Low Back Vowel Deletion (a-deletion)

$$\left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ +\text{bk} \\ -\text{long} \end{array} \right] \rightarrow \emptyset \text{ // } \left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \end{array} \right] \left[ \begin{array}{l} -\text{voc} \\ +\text{cons} \end{array} \right] \text{ — } \#$$

## CHAPTER 7

### ON SYLLABLE STRUCTURE

#### 7.1 GENERAL REMARKS

The relevance of syllable structure in any syllable timed language is an indisputable fact, be it Latin or Sanskrit or NIA languages. The various changes that such languages undergo during the course of their historical development almost invariably affect the syllable structure. The syllable structure is also relevant for stress or accent in such languages, as the latter depends on the former, i.e., on whether the syllable is heavy or light; stress is never an independent variable in such languages. Changes in syllable structures are directed towards the realisation of the preferred syllable structure of the language concerned. Schane 1972 has shown that in French there exist rules whose motivation is largely to yield the preferred syllable structure, which is  $C\bar{V}$  or CVC in French. Schane rightly points out that CV is the unmarked syllable structure according to the theory of markedness as developed by Chomsky and Halle 1968. However, we think that he is wrong in denying this status to a CVC syllable.

According to the SPE, given a sequence of two segments the first is unmarked for a consonant and the second for a vowel. This gives us CV, and the preponderance of CV syllables in the languages of the world proves the marking conventions to be correct; in fact, there are languages like Hawaiian, which have only the CV type. But if we have a sequence of three segments, then the third one is again unmarked for a consonant, and this gives us CVC. Thus both CV and CVC turn out to be unmarked syllable structures according to the theory. And the vast majority of Hindi syllables are either  $C\bar{V}$  or CVC. Incidentally, the motivation for the Sanskrit sandhi rule according to which two similar vowels coalesce into the corresponding long vowel can be given in terms of markedness of segment sequences, especially when the two relevant vowels are the even and odd segments in a sequence, counting from the left. In such a case we are left with the even segment only, the unmarked position for a vowel. The same motivation would be there for the Mid Vowel Rule (MVR(R); TR4', section 5.2.7) as well as for the Glide Rule (NRL, section 3.1). The other vowel contraction rules, e.g., the Low Vowel Rule (WH)(LVR(WH); TR5, section 5.2.6.4), can also be motivated in the same way. This might also explain why given MVR(R) and the Diphthong Rule (DIPH; PR16, section 5.2.6.4) in their grammars, some NIA languages or dialects have dropped



DIPH rather than MVR(R).

## 7.2 RULES AND THE PREFERRED SYLLABLE STRUCTURE

We have shown above that according to the theory of markedness of segment sequences, both CV and CVC are unmarked syllable types; we have also claimed that these two are the preferred syllable types in Hindi. If this is so, we should expect to find some rules of Hindi phonology which yield these syllable types, especially as Schane (1972) has shown this to be the case in French. And when we look at the rules we have formulated in the earlier chapters, we find that this is precisely what a number of rules do. That is, yielding an unmarked syllable type is one of their motivations. In this section, we shall show this to be the case with a number of rules that we have formulated. We shall first take up the rules which yield a C $\bar{V}$  structure and then those that yield a CVC structure.

### 7.2.1 Rules Yielding CV

#### 1. Compensatory Vowel Lengthening (CVL; TR2, 4.1.1.1)

This rule, operating on a sequence of the type VCCV, yields the type C $\bar{V}$ , e.g., given addhā 'half (m,sg)', ḍāṇḍā 'stick', etc., the rule gives ādḥā, ḍārī respectively.

#### 2. Final Vowel Lengthening (FVL; PR8, 4.1.1.2)

This rule by lengthening the final short high vowels of Sanskrit words, yields a final C $\bar{V}$  syllable. The

rule changes a CV syllable into a C $\bar{V}$  one, the motivation for this is the constraint in Hindi that tolerates only long vowels word finally.

### 3. Glide Rule (GR; NR1, 3.1)

By turning a high vowel immediately followed by another vowel into a glide, this rule yields a CV structure. This rule was first proposed by Schane (1972) as one of his natural rules.

### 4. MVR(R) (TR4', 5.2.7)

This rule by coalescing a and a following high vowel into the corresponding long mid vowel yields a  $\bar{V}$ ; if the preceding segment is a consonant, we get C $\bar{V}$ , e.g., /baik/ 'sell', /rauk/ 'run' become [bēc̣] and [rōk] respectively. (The consonantal alternation is irrelevant here.)

### 5. LVR(WH) (TR5, 5.2.6.4)

This rule, which exists in WH in place of Hindi DIPH, coalesces a and a following high vowel into the corresponding low vowel; if the preceding segment is a consonant, then in this case also we get C $\bar{V}$ .

## 7.2.2 a-Deletion

In contrast to the rules mentioned above, this rule yields a CVC structure, both medially as well as finally. It applies to forms like kamala 'lotus', kamalō̃ 'lotuses

(obl, pl)', phisalana 'slipperiness', etc. giving kamal, kamlō, phislan respectively.

All this is fairly straightforward and is, in a way, a recapitulation of what we have said earlier while formulating some of these rules. We now propose to examine the Stem Vowel Shortening Rule (SVS; PR9, 4.1.2.1), which affects syllable structure in a different way in the sense that instead of changing a segment sequence into a CV or CVC type, it affects the quantity of the syllable.

### 7.3 SYLLABLE QUANTITY

Kelkar (1968:26) has classified syllables in Hindi in the following way quantitatively.

1. Light: a syllable ending in a short vowel.
2. Medium : a syllable ending in a long vowel or having a short vowel checked by a single nonsyllabic.
3. Heavy : any other syllable.

Note that the onset of the syllable plays no role in determining the quantity of the syllable, although the coda does. We shall assign one mora to a light syllable, two to a medium one, and three to a heavy one.

As SVS simply shortens the stem vowel when a derivational suffix is added to the stem, noun or verb, it does not change the number of segments in the word. But clearly

it changes the number of moras of the word. More precisely, if the number of moras is going to be  $n$ , the application of SVS makes it  $n-1$ . It follows from this that the motivation of the rule is due to some constraint governing the number of moras in Hindi words. We might also infer from this why SVS is not applicable in inflection. Addition of inflectional endings does not result in the creation of a new word, but the addition of derivational suffixes most obviously does. As the constraint applies to derived words only, SVS applies in word derivation only. However, it can only be partly an explanation, as the shortening of final vowels does take place in plural inflection, although only high vowels are affected (PR11, 4.1.2.4). But even SVS at times yields a preferred syllable type in Hindi. Take, for example, words like ā̃m 'mango', bāṭ 'road', from which we get derived words like ā̃mrāī 'mango orchard', baṭmār 'highwayman'. The stems are heavy syllables which are turned into medium in the derived words. We have said that both  $C\bar{V}$  and  $CVC$  are preferred syllable types in Hindi, and both are medium. However, in most cases like this, a-deletion also plays a part.

#### 7.4 THE STRUCTURE OF WORD MEDIAL SYLLABLES

##### 7.4.1 Consonant Clusters in MIA and NIA

The simplification of OIA consonant clusters in MIA and NIA through assimilation, aspiration, affrication,

compensatory vowel lengthening etc. is a well accepted fact; consider, for example, Skt. hastin > Hindi hāthī 'elephant', Skt. sāṇḍhyā > H. sājh 'evening', Skt. adya > H. āj 'today', Skt. satya > H. sach 'truth', etc. Consequently, MIA and NIA should not have any consonant clusters left. This is true of the tadbhava part of the vocabulary, but with an important qualification: in both MIA and NIA it is the heterogeneous clusters that have been eliminated; geminate clusters, i.e.,  $-C_1C_1-$ , as well as the sequence of a nasal homorganic with the following consonant, i.e.,  $-NC-$ , are perfectly well tolerated; witness, e.g., Hindi words like kuttā 'dog', billī 'cat', gaḍḍhā 'pit', kāṅghā 'comb (n)', āṇḍar 'inside', āmbār 'pile (n)' and many others. In both these cases the two consonants are separated by a syllable boundary, i.e., kut=tā, ān=dar etc. (We represent a syllable boundary with =.) In other words, the first consonant constitutes the coda of the preceding syllable and the second consonant constitutes the onset of the succeeding syllable.

#### 7.4.2 Markedness and Consonant Clusters in Hindi

A consonant cluster is a marked segment sequence and, therefore, any reduction of clusters can always be motivated on the grounds of getting an unmarked segment sequence. Then, given this tendency towards simplification of consonant clusters in Hindi and other NIA languages, we would

not expect a complication of word structure and, consequently, of syllable structure. That is, given -CC-, -C- will be expected and desirable, but not -CCC-. NIA or no NIA, -CCC- is a much more marked segment sequence than -CC-. However, this is precisely what we get in Hindi in certain cases, i.e., a -CC- cluster is realised as a -CCC- cluster, and as far as we know, this is true of all Indian languages, IA or Dravidian. Consider, for example, Sanskrit words like anyāya 'injustice', satya 'truth', vidyā 'learning', vidvāna 'scholar', putra 'son', bhadra 'cultured', ākramaṇa 'invasion', vākya 'sentence (n)', āmra 'mango', etc. These words are pronounced in Hindi as [ˈannyāy], [sattya], [viddyā], [viddvān], [puttra], [bhaddra], [ākkrāmān], [vākkya] and [āmmra] respectively. [We shall assume for the present that in Sanskrit these words were pronounced as they have been given earlier. We shall return to this question later.] What is happening here is clear; in the place of two medial consonants we get three. That is, an already marked sequence is being replaced by a still more marked one. That this process belongs to the core of Hindi phonology is clear from the fact that English and Arabo-Persian words of the same type are also pronounced with an extra medial consonant. Thus, English words like declaim, depression, afraid, and Arabo-Persian words like umra 'age (n)', hijra 'separation', sabra 'patience', etc.

are pronounced as [dikklēm], [dippreśan], [affrēd], [ummra], [hi<sup>ʃ</sup>j<sup>ʃ</sup>ra] [sabbra] respectively. Not only this. In Hindi there is a metathesis rule which operates on sequences like -hm-, -hn-, -hl-, -hv- and changes them to -mh-, -nh-, -lh- and -vh- respectively; after this the consonants are naturally aspirated (TRL, 2.1). But the interesting thing is that the consonants are not just aspirated; they are geminated too. Thus, words like āhlāda 'joy', brahma 'Brahma', vahni 'fire', jihvā 'tongue' are pronounced as [āl<sup>h</sup>hād], [brammha], [vannhi], [jivv<sup>h</sup>ā] respectively. This medial consonant gemination does not occur in all cases; on the basis of the data given, it takes place only when the first segment is a stop or a nasal and the second one is r or y/v. Thus, there is no gemination in the case of words like bhasma 'ashes', ratna 'jewel', ambar 'sky'; it also does not take place if there are already three consonants in the medial position, as in chandra 'moon', etc.

Although the rule for medial consonant gemination can be easily formulated, we do not think that the main question here is one of writing a rule. The main question here is rather one of finding a motivation, if possible, for this complication of syllable structure in the word medial position. Because if we have three consonants in the medial position, then either the coda of the preceding syllable or the onset of the succeeding one is going to have two consonants; of course, it

is possible for either to have all the three, but this is hardly any better. In any case the syllable structure becomes complicated.

#### 7.4.3 Motivation for Medial Consonant Gemination

Such a motivation does exist in our opinion. In the beginning of this section we said that the only consonant clusters that are really native in Hindi are of the type  $-C_1C_1-$  and  $-NC-$ , where N is homorganic with C. We are not talking of the phonetic level here; phonetically, because of the deletion of a medial a, there are many more types of clusters. We are also not talking of the sequence Ch from which we derive C<sup>h</sup>, as neither is relevant here. We have also stated that in both the cases the syllable boundary separates the two consonants, i.e., the first consonant belongs to the coda of the preceding syllable and the second one to the onset of the succeeding one. Thus words like billī 'cat' and bandar 'monkey' can be represented as bil=llī and ban=dar. The difference between a homorganic nasal and the stop is not very great; if the stop is voiced, the difference consists in the specification for the features [son] and [nas]; if the stop is voiceless, the difference consists in the specification for the feature [voi] as well. Thus, both  $C_1C_1$  and NC can be subsumed under  $C_1C_1$ , more precisely,  $XC_1=C_1Y$ . Now consider



a word like putra 'son' in this light. The division here is put=ra, or  $XC_1=C_2Y$ , but this goes against the trend of configuration that we get in native words; in this case, the coda of the first syllable and the onset of the second are not the same. We suggest that medial consonant gemination is one way of resolving this contradiction. After medial consonant gemination, we get put=tra, or  $XC_1=C_1C_2Y$ . The inapplicability of this rule in the case of words like ratna 'jewel', bhasma 'ashes' is due to another constraint on syllable structure. If the rule applies to ratna, we shall get rat=tna, but a stop+nasal sequence cannot constitute the onset of a syllable. This constraint is shared by Sanskrit as well as many other Indo European languages, if not by all of them; in Sanskrit the only exception is jñ, as in jñāna 'knowledge', and in one case, ghn, as in ghnanti 'they kill'. If the rule applies to bhasma, it will give us bhas=sma. Now, although sm is all right as a syllable onset in Sanskrit, English and many other languages, it is not so in Hindi. For example, given  $\neq sC$ , an epenthetic i is inserted before s by Hindi speakers in both Sanskrit and English words. This epenthesis takes place in other NIA languages also, although in some cases the epenthetic vowel is a. It is evident that after epenthesis yields isC, we get the same type of structure that we already have in bhasma, i.e.,

a vowel + s + another consonant. Thus, when words like smāarak 'memorial', ślōk 'verse', snān 'a bath' become [ismāarak], [iślōk], [isnān] respectively after epenthesis, in terms of syllable division we get is=māarak, iś=lōk, and is=nān. But s=m is already there in bhasma. On the basis of further data from the language, we can say that  $C_1=C_2$  is tolerated if:

1.  $C_1$  is a fricative, i.e., s, ś, ṣ, and  $C_2$  is a consonant, i.e., a stop or a fricative or a nasal. Thus, there is no medial gemination in the case of words like bhasma 'ashes', usturā 'razor blade', kaṣṭa 'pain (n)', susta 'slow', niščal 'immobile', etc.  
Or
2.  $C_1$  is a nasal and  $C_2$  is a consonant, e.g., words like andar 'inside', ambar 'sky', kanghā 'comb (n)', etc.  
Or
3.  $C_1$  and  $C_2$  both are consonants, i.e., stops or fricatives or nasals. For example, there is no gemination in words like rakta 'blood', apsarā 'celestial nymph', ratna 'jewel', yatna 'effort', etc. Or
4.  $C_1$  is a liquid, i.e., r or l, and  $C_2$  is a consonant. For example, in words like darpan 'mirror', kalpanā 'imagination', arka 'essence', pragalbha 'talkative', etc., there is no gemination.

It will be clear from an examination of these four exceptions to medial consonant gemination that the controlling factor is the nature of  $C_2$ ;  $C_2$  has to be either a glide, more precisely, a high glide, or a liquid. If  $C_2$  is an obstruent or a nasal, no gemination can take place. We can state the rule as follows.

PR19: Medial Consonant Gemination

$$\emptyset \rightarrow [-\text{syll}]_i / \left[ \begin{array}{c} +\text{voc} \\ -\text{cons} \end{array} \right] [-\text{syll}]_i = \left[ \begin{array}{c} +\text{voc} \\ \alpha\text{cons} \\ -\alpha\text{hi} \end{array} \right] \left[ \begin{array}{c} +\text{voc} \\ -\text{cons} \end{array} \right] \left[ +\text{native} \right]$$

In Hindi, as it is only vowels which are syllabic,  $[-\text{syll}]$  includes all non-vowel segments, i.e., pure consonants, glides and liquids;  $\left[ \begin{array}{c} +\text{voc} \\ \alpha\text{cons} \\ -\alpha\text{hi} \end{array} \right]$  refers to either the high glides, i.e.,  $\underline{y}$  and  $\underline{v}$ , or the liquids, i.e.,  $\underline{l}$  or  $\underline{r}$ . As  $\underline{y}$  and  $\underline{v}$  are  $[+\text{hi}]$  but  $\underline{l}$  and  $\underline{r}$  are  $[-\text{hi}]$ , the specification  $[-\alpha\text{hi}]$  is required. The specification  $[+\text{native}]$  ensures that the rule will apply when the word in question is being pronounced in the normal way; hence, the possibility of pronouncing depress, declare, etc. without the medial gemination always exists. Maddah (1972) has said that medial consonant gemination does not take place in Urdu pronunciation of Arabo-Persian words, although it applies to words of Sanskrit and English origins. That is, in the case of Urdu, the specification will be  $[-\text{Arabo-Persian}]$ .

We would like to add that, as far as we know, Maddah (1972:2) is the only writer who has mentioned medial consonant gemination specifically for an NIA language.

#### 7.4.4 Medial Consonant Gemination and Sanskrit

The rule for medial consonant gemination that we have given above (PR19) will account for the pronunciation of all those words which have a medial  $-C_1C_2-$  sequence, where  $C_1$  is not a nasal homorganic with  $C_2$ . Furthermore, we have done so purely on the basis of synchronic data. In this section, we would like to examine whether this rule existed in Sanskrit or not; note that in 7.4.2 we assumed that there was no medial gemination in Sanskrit. If this rule existed in Sanskrit, then it would mean that Hindi and other NIA languages have not added a new rule to their grammars, but have retained an old one. It would also account for medial gemination found in the pronunciation of relevant Sanskrit words in Dravidian languages; they have simply taken the pronunciation of Sanskrit speakers of the time when these words were borrowed into Dravidian languages. Second, if this rule existed in Sanskrit, it would be interesting to examine what changes, if any, have taken place between the Sanskrit rule and the Hindi rule given above. (In fact, our rule belongs not only to a Pan-NIA grammar but rather to a Pan-Indian one.)

For evidence regarding the existence of this rule in Sanskrit, we have based our remarks on Varma 1961. (The book was first published in 1933.) According to Varma, medial gemination, or the doubling of consonants as he calls it, goes back to Vedic as is evidenced by the rules regarding it in the various Prātishākhyaś and Shikṣāś; it is also mentioned by Pāṇini and Patanjali. A number of Vedic manuscripts indicate this doubling in writing also, e.g., addya for adya 'today', puttrēṇa for putrēṇa 'son, inst sg', etc. (p. 63). In Gupta inscriptions (c.5-6c. A.D.) we get the same thing, e.g., puttrasya for putrasya 'son, gen sg', dīrggha- for dīrggha- 'long (adj)', etc. (p.64); note the non-gemination of s in the former. It should be mentioned here that there is some laxity regarding the indication of the doubling in orthography. For example, some Vedic manuscripts indicate doubling in -Cy- or -Cy- sequences whereas some others do not. Gupta inscriptions, as a rule, do not indicate doubling in -Cy- or -Cy- sequences. Then, the doubling of s is never indicated, although that of ṣ and ś is. We are inclined to agree with Varma when he says that the non-doubling of s was perhaps just an orthographic convention and in actual pronunciation s was regularly doubled in relevant contexts. We can compare this convention with the modern convention followed in writing Hindi and other Indian languages. In

Hindi, medial consonant gemination is never indicated in writing, although in pronunciation it is very much present.

But, although medial consonant gemination exists in both Sanskrit and Hindi, syllabic division is different in the two languages. Given  $-C_1C_1C_2-$ , in Hindi the division is always  $-C_1=C_1C_2-$ ; that is, the first consonant constitutes the coda of the preceding syllable and the second and the third consonants constitute the onset of the succeeding one. This was allowed in Sanskrit as well. But in Sanskrit a second option was also available, i.e.,  $-C_1C_1=C_2-$ ; in this case the coda of the preceding syllable has two consonants whereas the onset of the succeeding syllable has only one consonant. Needless to say, the two divisions are minor images of each other. Thus we find that although the rule has not changed during the development of IA languages, syllable division has changed in the sense that it has become more restricted in Hindi than it was in Sanskrit.

## CHAPTER 8

### CONCLUSION

#### 8.1 THE AIM OF THE DISSERTATION

We undertook this study with a view to examining some aspects of Hindi phonology which seemed to us interesting and relevant not because they had not been discussed by earlier writers in the field but because they offered a challenge in the form of possible alternative analyses. In fact, most of the topics that we have discussed in this dissertation have been dealt with by earlier linguists following either the TG framework or the neo-Bloomfieldian one; in some cases the topics have been dealt with by the traditional grammarians as well. For example, most of the traditional grammarians like, for example, Guru 1963, Vajpayee 1977, Kellogg 1955, Sharma 1975, Beames 1966, etc., have something to say about the qualitative and quantitative vowel alternations and a-deletion.

#### 8.2 MAIN THEMES

Apart from the analyses of the problems that we were concerned with in this dissertation, we think that we have been able to isolate certain issues of some importance to

both Hindi as well as NIA linguistics which are more important than the specific solutions proposed by us. These issues can be stated as follows:

1. The relevance of historical and/or dialectal information to a synchronic description.
2. The abstractness of underlying representations.
3. Two sources for a segment vs. only one.
4. Inflectional vs. derivational morphology.
5. The relevance of grammatical information in phonology.
6. The importance of syllable structure.
7. Quantitative vs. qualitative vowel alternations.

All these issues have come up during our discussion of the various problems in the main body of the dissertation; as such they have been discussed there. In what follows, we shall make only brief comments on each one of these in order to present an overall view. We shall take these up in the order given.

### 8.2.1 The Relevance of Historical and Dialectal Information

#### 8.2.1.1 The relevance of historical information

Wherever we have referred to diachronic information, e.g., in the section on aspiration (2.1) and in our discussion of the underlying representation of mid vowels (Chapter 5), we have made it quite clear that facts from the history of the language under study can only be used as supporting



### 8.2.1.2 The relevance of dialectal information

Much of what we have said regarding historical information will apply in this case also. We do not confuse between the grammar of a language and the comparative grammar of languages related to it, or of its dialects. But if the grammar of a language apart from describing and explaining facts about that language can do so for its dialects as well, it is definitely a better grammar than one which does not do so. As an example, we can take Kellogg. His grammar was first published in 1875, but even after more than a hundred years, it remains one of the best grammars of Hindi. But throughout his book, and especially in the nominal and verbal paradigms, Kellogg has given examples from other languages related to Hindi. However, he has neither confused Hindi with those languages nor the other way round; it is because of this that he remains one of the best.

Bringing in information from dialects and/or related languages, i.e., being aware of the dialect situation, can help us in being more objective in our statements. For example, while discussing vowel alternations in Chapter 4 (4.1) as well as in our discussion of Hindi mid vowels (5.2), we mentioned that in a large number of cases a short mid vowel varies freely with its short high counterpart. (This is true of the author's speech.) We also mentioned that

Standard Hindi as well as Urdu prefer the short high vowel, and that the short mid vowel is due to dialect borrowing from Avadhi and other NIA languages spoken in the Hindi speaking areas where we get the short mid vowel. Now if we had not mentioned this fact about the dialects, we would have had to say that the mid vowel in such cases is non-standard, an unnecessary value judgement. But attributing the variation to dialect borrowing has no such pejorative connotation. Of course, we could have always taken "our dialect" as the standard as is the common practice; either way the statement would not have been truly objective. We have throughout tried to avoid the terms "Standard Hindi" or the "standard dialect" as far as possible simply because we did not want to make unnecessary value judgements. We have avoided the term "our dialect" in the sense of the "standard dialect" because after Postal's equating "our dialect" with Standard American English in his 'Remind' paper, we have become wary of the term.

### 8.2.2 The Abstractness of Underlying Representations

Our position in this regard has been that of Kiparsky 1968b, 1971, Wang 1971, Vachek 1970 etc. We believe that in the absence of surface alternations, the phonetic form should be taken as the underlying form, and that absolute neutralization is a very costly feature of grammar. But

we do not think that these should be taken as inviolable tenets, nor do we think that abstractness should be pursued for its own sake. Our own analysis of mid vowels can be said to be abstract in the sense that the underlying segments, i.e., ai for ē and au for ō, never surface as such. However, we have justified our analysis on the ground that it is able to explain certain facts about Hindi and other NIA languages in a very simple way, something that is not possible if we posit the mid vowels underlyingly or derive them from underlying short high vowels. In fact, we have shown (5.2.4) that this latter analysis results in a very costly grammar indeed. In the final analysis what counts is the explanatory power of the solution given. But we do not think that the same can be said about Chomsky and Halle's positing [x] in the underlying representation of words like right; we agree with Kiparsky that this amounts to the diacritic use of a phonological feature (Kiparsky 1968b). To account for the similarity between right/righteous and Christ/Christianity, we do not need [x] in right, parallel to [s] in Christ; we need marking.

In a similar manner we would defend deriving C<sup>h</sup> from Ch (2.1), of deriving high glides from high vowels followed by another vowel (3.1), and positing a word final a for all words which phonetically end in a consonant (Chapter 6). We

have shown that whenever an unaspirated consonant and h come together, the former gets aspirated and the latter is deleted. We have also shown in 2.1 that the cluster approach to aspiration is superior to a unit segment approach as it is much closer to reality both synchronically and diachronically. Regarding the word final a, we have given a number of contexts where it does surface; we have also shown that if we do not posit a word final a, we will not be able to account for flaps in a natural way. We will also not be able to account for a number of common Sanskrit compounds without the help of quite a few ad hoc and otherwise totally unnecessary rules. As far as the high glides are concerned, we would like to propose one very good test of the reality of this rule. Whenever a Hindi speaker has to pronounce a high vowel followed by another vowel, the high vowel inevitably turns into a glide; this is precisely what our rule (NRL), or rather Schane's rule, says. Consequently, none of our analyses can be really called abstract.

### 8.2.3 Two Sources for a Segment vs. only One

In our opinion the genesis of this problem goes back to Chao's classic 'Non-uniqueness' paper published in 1934. In this paper Chao pointed out that for a given set of phonetic data it is possible to have more than one phonemic solution, i.e., an indeterminacy can exist. Subsequent

neo-Bloomfieldian linguists have been at pains to achieve a unique phonemic solution; the principle of once-a-phoneme-always-a-phoneme is a direct result of this search for uniqueness. Unfortunately, this requirement of a unique phonemic solution has become as much a part of generative phonology as it was of Neo-Bloomfieldian linguistics; the only difference is that in generative phonology it is called rule optimization. Schane (1968a) has proposed that if a segment can be derived from two sources, then the unmarked one of the two should be chosen. Again, it is linguists believing in a concrete approach to phonology who have accepted non-uniqueness of phonological representations, e.g., Kiparsky 1968b, Vachek 1970, Malone 1970, Wang 1971, Shibatani 1971. We have given both unique and non-unique solutions in this work. For example, we have derived all aspirated consonants from an underlying sequence of the corresponding unaspirated consonant and h; in support of this we have shown that this analysis is superior to one which derives some aspirated consonants from a sequence like the one mentioned above and posits some other aspirated consonants underlyingly. But in the case of nasalised vowels and diphthongs we have opted for a non-unique solution. A majority of nasalised vowels can be derived through the Vowel Nasalization Rule (VNR; PRL2, 4.2.2.1) and we have derived them so. But there are quite a few cases in

which nasalised vowels cannot be so derived without undue complication of the grammar; for this reason we have posited them underlyingly as [+nas] (see 4.2). Likewise, we have derived diphthongs also from two sources. Some diphthongs come from underlying ai and au, whereas some others come from underlying aia and aua after the deletion of the second a by the rule deleting a post-glide a (PR6, Section 3.3). While discussing the diphthong rule (PR16, Section 5.2.6.4) we have given reasons for our decision to do so. In neither of these two cases will a unique solution be sufficiently motivated, whereas in the case of aspiration it is. We feel that it is only in the light of specific problems of a language that we can decide the question of non-uniqueness and uniqueness, and that neither can be accepted as an inviolable principle of linguistic analysis.

#### 8.2.4 Inflectional vs. Derivational Morphology

Our work on vowel alternations related to the nominal as well as verbal paradigms has shown that a very clear distinction between inflectional and derivational morphology exists in Hindi. For example, the shortening of the stem vowel, either in a verb or a substantive, takes place only when a derivational suffix is added to the stem (4.1.2.1). This vowel shortening applies to all vowels and to all cases of word derivation, non-compounds as well as tadbhava compounds.

On the other hand, the only instance of vowel alternation in inflectional morphology is the shortening of a phonetically long final high vowel when the plural suffix is added to the word. As this shortening affects only final high vowels, and that too only in the case of plurals, it cannot be said to have anything in common with the shortening of the stem vowel mentioned earlier. Therefore, while we can say that the shortening of the stem vowel is a characteristic feature of word derivation, we cannot make the same claim for inflection. As in most cases the inflectional suffixes, nominal or verbal, consist of one syllable whereas the derivational ones consist of two, it is quite tempting to attempt to relate it to some kind of constraint regarding the number of syllables in a Hindi word. However, this is not possible as the future tense forms in Hindi always have two syllables after the verbal stem, e.g., jā-ē-gā 'he will go' has two syllables after the stem jā 'go'. Perhaps the motivation of stem vowel shortening lies in making a distinction between the phonetic shape of a word when it is used as an independent word and when it is used as part of a derived word. If this explanation is correct, it also explains why there is no shortening of the stem vowel in inflection; an inflected word is still the same as the uninflected one intuitively, but a derived word is not.

That is, boy and boys are much closer to each other than boy and boyish.

#### 8.2.5 The Relevance of Grammatical Information

The idea of using grammatical information in phonemic analysis was anathema to Neo-Bloomfieldian linguists, Pike being the sole exception (Pike 1947, 1952). This was a legacy of the Neo-Grammarian tradition, which maintained that grammatical information was irrelevant to sound change. Although the Neo-Grammarian position was challenged by dialectologists and linguistic geographers like Schuchardt and Gilliéron, by and large it was accepted by most linguists. In fact, the conception of an autonomous phonemics is simply an extension of the doctrine from historical linguistics to descriptive linguistics.

Generative phonology has never accepted the non-relevance of grammatical information to phonology. Our position in this regard is that phonological rules may refer to grammatical information, and our work has amply illustrated this. There are quite a few rules which refer to a morpheme boundary, or a word boundary, or a suffix; all these are grammatical categories. For example, high vowels are lengthened word finally (4.1.1.2); flapping of intervocalic voiced retroflex stops does not take place if there is a formative boundary between the first vowel and the stop (6.2.1);



a stem vowel is shortened when a derivational suffix is added to the stem (4.1.2.1), etc. Ohala's rule (6.2) also makes reference to a formative boundary. On the other hand, rules like the Aspiration Rule (TR1, section 2.1), the Glide Rule (NR1, section 3.1), Post-Glide a-deletion (PR6, section 3.3), the Vowel Nasalisation Rule (PR12, section 4.2.2.1), etc. do not require any grammatical information.

### 8.2.6 The Importance of Syllable Structure

That syllable structure should be important in a syllable-timed language is evident, and Hindi is no exception. Just as Schane (1972) has done with regard to French, we also have been able to show that there exist a number of rules in Hindi which are aimed at the realisation of the preferred syllable structure, i.e., either  $C\bar{V}$  or  $CVC$  (7.2). Both of these are unmarked syllables; as such, rules resulting in these structures are naturally to be expected. However, what is much more important in this connection is the fact that a rule resulting in a highly marked segment sequence of  $-CCC-$  in the word medial position (PR19, section 7.4.3) can be motivated only on the basis of syllable structure. To be more precise, it is the syllable division permitted word medially which is the guiding factor behind the rule. As  $C_1=C_1$  is very well tolerated in Hindi, as well as in other Indian languages,  $C_1=C_2$  is

converted into  $C_1=C_1C_2$  in a number of cases by medial consonant gemination.

### 8.2.7 Quantitative vs. Qualitative Vowel Alternations

We have shown in 4.1 that Hindi shows only quantitative vowel alternations and that the apparent cases of qualitative alternations can be reduced to quantitative ones. This is quite a contrast to Sanskrit where vowel alternations are both qualitative and quantitative; in fact, the former constitute a majority as both the high vowels show it. Most of the quantitative alternations are due to the vriddhi of a. This, then is a striking difference between the two languages, and it would be interesting to examine the situation in other NIA languages. Perhaps the distinction is related to syllable quantity. Thus, when a vowel is lengthened or shortened, the syllable quantity changes; no such change need be there in a qualitative alternation. For example, if a long vowel becomes a long mid one or vice versa, syllable quantity does not change. If so, the vowel alternations in Hindi can be related to syllable structure in terms of syllable quantity.

### 8.3 CONCLUDING REMARKS

In the preceding section we have outlined what seemed to us some important points which emerge from this dissertation. Needless to say that we do not consider this list to be

exhaustive. However, the fact that these points have kept recurring shows that they are fairly significant. Although we have not considered data from other NIA languages in relation to all of these, we feel that the point mentioned above will be important in their case also. As far as points like the abstractness of underlying representations, the relevance of historical, dialectal and grammatical information, non-uniqueness of underlying representations are concerned, these are theoretical issues and, as such, will be relevant to phonological descriptions in general. On the other hand, as far as the points relating to quantitative vs. qualitative vowel alternations, inflectional vs. derivational morphology, and the importance of syllable structure are concerned, they will be of greater relevance to NIA studies, although the relevance of these to Dravidian language studies also seems apparent. It was not our aim to deal with Pan-NIA grammar, and we have not done so. However, some of our rules seem to be definitely Pan-NIA; for example, the Mid Vowel Rule (TR4', 5.2.7) can be said to exist in the grammars of at least Hindi, Gujarati, Rajasthani and Avadhi; medial consonant gemination (PR19, 7.4.3) takes place in all Indian languages. As far as the deletion of a, more precisely, of a short low back vowel, is concerned, it again seems to exist in all NIA languages.

(We are referring here to the deletion of a in both the medial and final positions.)

Consequently, we feel that if future work in NIA languages examines other NIA languages in relation to the points that we have mentioned here in connection with Hindi, we can get much closer to a Pan-NIA grammar. The close similarity of various NIA languages to each other is generally recognised. But it is only a Pan-NIA grammar which will be able to show this similarity, as well as the difference, in an explicit way.

## APPENDIX

We give below, in their final form, the rules which we have formulated in the various chapters. The rules which we formulated but rejected later are not included.

### CHAPTER 2

(1) TR1: Aspiration (p.16)

$$\begin{array}{ccc} \text{C} & \text{h} & \\ 1 & 2 & \end{array} \rightarrow \begin{array}{cc} 1 & 2 \\ [+asp] & \emptyset \end{array}$$

(2) PR1: a- to -e Rule (p.34)

$$\left[ \begin{array}{c} +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ -\text{long} \end{array} \right] \rightarrow \left[ \begin{array}{c} -\text{lo} \\ -\text{bk} \end{array} \right] / \left[ \begin{array}{c} +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ -\text{long} \end{array} \right] \text{h} - \left\{ \begin{array}{c} [+cons] \\ \# \end{array} \right\}$$

(3) PR2: u- to -o Rule (p.36)

$$\left[ \begin{array}{c} +\text{voc} \\ -\text{cons} \\ +\text{hi} \\ +\text{bk} \\ -\text{long} \end{array} \right] \rightarrow [-\text{hi}] / \left[ \begin{array}{c} +\text{voc} \\ -\text{cons} \\ +\text{lo} \end{array} \right] \text{h} \left[ \begin{array}{c} +\text{cons} \\ \end{array} \right] \left[ +\text{A/P} \right]$$

(4) PR3: i- to -e Rule (p.40)

$$\left[ \begin{array}{c} +\text{voc} \\ -\text{cons} \\ +\text{hi} \\ -\text{bk} \\ -\text{long} \end{array} \right] \rightarrow [-\text{hi}] / \left[ \begin{array}{c} +\text{voc} \\ -\text{cons} \\ +\text{lo} \end{array} \right] \text{h} \left[ \begin{array}{c} +\text{cons} \\ \end{array} \right] \left[ +\text{A/P} \right]$$

[A/P = Arabo-Persian]

CHAPTER 3

(5) NR1: Glide Rule (GR) (p.57)

$$\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{hi} \end{bmatrix} \rightarrow [-\text{voc}] / \text{---} \begin{bmatrix} +\text{voc} \\ -\text{cons} \end{bmatrix}$$

(6) PR4: a- Raising (p.61)

$$\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{lo} \end{bmatrix} \rightarrow \begin{bmatrix} \alpha\text{bk} \\ -\text{lo} \end{bmatrix} / \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{hi} \\ \alpha\text{bk} \end{bmatrix} + \text{---} + \begin{bmatrix} -\text{voc} \\ -\text{cons} \\ +\text{lo} \end{bmatrix}$$

(7) PR5" : Glide Deletion (p.66)

$$\begin{bmatrix} -\text{voc} \\ -\text{cons} \\ +\text{hi} \\ \alpha\text{bk} \end{bmatrix} \rightarrow \emptyset / \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{hi} \\ \alpha\text{bk} \end{bmatrix}$$

(8) PR6: Post-Glide a- Deletion (p.66)

$$\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ -\text{long} \end{bmatrix} \rightarrow \emptyset / \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{lo} \end{bmatrix} \begin{bmatrix} -\text{voc} \\ -\text{cons} \\ +\text{hi} \end{bmatrix} \text{---}$$

(9) PR7: Glide Lowering (GL) (p.66)

$$\begin{bmatrix} -\text{voc} \\ -\text{cons} \end{bmatrix} \rightarrow [-\text{hi}] / \begin{bmatrix} +\text{voc} \\ -\text{cons} \end{bmatrix} \text{---} \left\{ \begin{bmatrix} +\text{cons} \\ \# \end{bmatrix} \right\}$$



CHAPTER 5

(15) PR15: High Vowel Lowering (HVL) (p. 128)

$$\left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{hi} \\ -\text{long} \end{array} \right] \rightarrow \left[ \begin{array}{l} -\text{hi} \\ +\text{long} \end{array} \right] / X - Y \left\{ \begin{array}{l} [+C1 \text{ II}] \\ S \\ [+stem] \end{array} \right\}$$

(16) PR10': Mid Vowel Raising Rule (R) (p.139)

$$\left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ -\text{hi} \\ -\text{lo} \\ -\text{long} \end{array} \right] \rightarrow [+hi] / X - Y + \begin{array}{l} Z \\ [+der] \\ \text{suff} \end{array}$$

(17) PR16: Diphthong Rule (DIPH) (p. 151)

$$\left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{hi} \end{array} \right] \rightarrow [-\text{voc}] / \left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ +\text{bk} \end{array} \right] - \left\{ \begin{array}{l} [+cons] \\ \# \end{array} \right\}$$

(18) TR5: Low Vowel Rule (LVR (WH)) (p.151)

$$\begin{array}{c} \left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ +\text{bk} \\ -\text{long} \end{array} \right] \\ 1 \end{array} \quad \begin{array}{c} \left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{hi} \\ -\text{long} \\ \alpha\text{bk} \\ \alpha\text{rnd} \end{array} \right] \\ 2 \end{array} \rightarrow \begin{array}{c} \begin{array}{l} 1+2 \\ +\text{lo} \\ \alpha\text{bk} \\ \alpha\text{rnd} \end{array} \end{array}$$

(19) TR6: Vowel Metathesis (Rule for Avadhi) (p.157)

$$\begin{array}{c} \left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{bk} \\ +\text{lo} \\ +\text{long} \end{array} \right] \\ 1 \end{array} \quad \begin{array}{c} \left[ \begin{array}{l} +\text{voc} \\ -\text{cons} \\ +\text{hi} \end{array} \right] \\ 2 \end{array} \quad \begin{array}{c} X \\ -\text{der} \\ \text{suff} \end{array} \rightarrow \begin{array}{ccc} 2 & 1 & 3 \end{array}$$



(20) TR4': Mid Vowel Rule (MVR(R))(p.161)

$$\begin{array}{ccc}
 \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ +\text{bk} \end{bmatrix} & \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{hi} \\ \alpha\text{bk} \\ \alpha\text{rnd} \end{bmatrix} & \left\{ \begin{array}{c} [+ \text{cons}] \\ \# \end{array} \right\} \rightarrow \begin{array}{c} 1 + 2 \quad 3 \\ \begin{bmatrix} -\text{hi} \\ -\text{lo} \\ \alpha\text{bk} \\ \alpha\text{rnd} \\ +\text{long} \end{bmatrix} \end{array}
 \end{array}$$

## CHAPTER 6

(21) (20) -Deletion: Ohala's formulation (p.167)

$$\partial \rightarrow \emptyset / \text{VC}_1^2 \text{ --- CV} / \left\{ \begin{array}{l} [+ \text{loan} \\ [+ \text{casual speech}] \\ [+ \text{normal tempo}] \end{array} \right\}$$

(22) PR17: Flapping Rule (p.171)

$$\begin{bmatrix} +\text{obs} \\ +\text{ab rel} \\ +\text{voi} \\ +\text{ret} \end{bmatrix} \rightarrow [+ \text{flap}] / \begin{array}{c} \text{V} \quad \text{X} \quad - \quad \text{V} \\ \begin{bmatrix} -\text{seg} \\ -\text{WB} \\ -\text{FB} \end{bmatrix} \end{array}$$

(23) PR18: Short Low Back Vowel Deletion (a-deletion)(p.189)

$$\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{lo} \\ +\text{bk} \\ -\text{long} \end{bmatrix} \rightarrow \emptyset / \begin{bmatrix} +\text{voc} \\ -\text{cons} \end{bmatrix} \begin{bmatrix} -\text{voc} \\ +\text{cons} \end{bmatrix} \text{ --- } \#$$

## CHAPTER 7

(24) PR19: Medial Consonant Gemination (p.202)

$$\emptyset \rightarrow [-\text{syll}]_i / \begin{bmatrix} +\text{voc} \\ -\text{cons} \end{bmatrix} [-\text{syll}]_i = - \begin{bmatrix} \alpha\text{voc} \\ \alpha\text{cons} \\ -\alpha\text{hi} \end{bmatrix} \begin{bmatrix} +\text{voc} \\ -\text{cons} \end{bmatrix} [+ \text{native}]$$

## BIBLIOGRAPHY

- Allen, W.S. 1953. Phonetics in Ancient India, London Oriental Series 1, Oxford University Press.
- Anderson, Stephen R., and Kiparsky, Paul. 1972. A Festschrift for Morris Halle, Holt, Rinehart and Winston.
- Anderson, Stephen R. 1974. The Organization of Phonology, Academic Press, New York.
- Bach, Emmon. 1968. 'Two Proposals concerning the simplicity metric in phonology,' Glossa 2.128-149.
- Bach, Emmon, and Harms, Robert T. 1968. Universals in Linguistic Theory, Holt, Rinehart and Winston, Inc.
- Bahri, Hardev. 1966. Gramin Hindi Boliyan, Kitab Mahal, Allahabad.
- Beames, John. 1966 (1872-79). A Comparative Grammar of the Modern Aryan Languages, Munshi Ram Manohar Lal, Delhi.
- Bhasin, Sharda. 1978. 's-deletion : in reference to the sequence structure of Hindi,' Language Forum, 4.2.
- Bhatia, Tej, and Kenstowicz, Michael. 1972. 'Nasalization in Hindi: A reconsideration,' Papers in Linguistics 1. 248-282.
- Bloomfield, Leonard. 1963. Language, Motilal Banarsidass, Delhi.
- Brame, Michael K. 1972. 'On the Abstractness of phonology: Maltese S,' Contributions to Generative Phonology, 22-61, ed. by M. K. Brame, University of Texas Press, Austin.
- Brame, Michael K. (ed). 1972. Contributions to Generative Phonology, University of Texas Press, Austin.
- Burrow, Thomas. 1955. The Sanskrit Language, Faber and Faber, London.

- Cairns, Charles E. 1969. 'Markedness, neutralization, and Universal redundancy rules,' Language 45.IV 863-885.
- Chafe, Wallace L. 1968. 'The ordering of phonological rules,' IJAL 34.115-136.
- Chao, Yuen Ren. 1934. 'The non-uniqueness of phonemic solutions of phonetic systems', in Joos 1963, 38-54.
- Chatterji, Suniti K. 1970. The Origin and Development of the Bengali Language, 2 Vols., Allen and Unwin.
- Chomsky, Noam. 1964a. 'Current issues in linguistic theory', in Fodor and Katz 1964, 50-118.
- Chomsky, Noam. 1964b. 'Topics in the theory of generative grammar,' in Current Trends in Linguistics, Vol. III, 1-60, ed. by T. A. Sebeok, The Hague.
- Chomsky, Noam. 1965. Aspects of the Theory of Syntax, M.I.T. Press.
- Chomsky, Noam. 1967. 'The Nature of phonological rules,' Language 43.102-128. Also in Makkai 1972, 529-548.
- Chomsky, Noam. 1969. 'Deep structure, surface structure and semantic interpretation', reproduced by Indiana University Linguistic Club.
- Chomsky, Noam, and Halle, Morris. 1965. 'Some controversial questions in phonological theory,' JOL 1.97-138. Also in Makkai 1972, 457-485.
- Chomsky, Noam, and Halle, Morris. 1968. The Sound Pattern of English, New York, Harper and Row.
- Coats, H.S. 1970. 'Rule environment features in phonology,' Papers in Linguistics 2.110-140.
- Crothers, John. 1971. 'On the abstractness controversy', POLA Reports, second series, No. 12, CRI-CR29, Phonology Laboratory, University of California, Berkeley.
- Dingwall, William Orr (ed.). 1971. A Survey of Linguistic Science, University of Maryland.
- Dixit, Ram Avatar. 1963. The Phonemes of Standard Hindi, Unpublished M.A. dissertation, University of Texas, Austin.

- Edgerton, Franklin. 1943. 'The Indo-European semivowels', Language 19.II 83-124.
- Fodor, Jerrold A. and Katz, J.J. 1964. The Structure of Language: Readings in the Philosophy of Language, Prentice-Hall, Englewood Cliffs, New Jersey.
- Fries, Charles C. and Pike, Kenneth L. 1949. 'Coexistent phonemic systems,' Language 25. 29-50.
- Fudge, Erik C. (ed.). 1973. Phonology, Penguin Books, England.
- Guru, Kamta Prasad. 1963 (1921). Hindi Vyakarana, Nagari Pracharani Sabha, Kashi.
- Halle, Morris. 1962. 'Phonology in generative grammar', Word 18.54-72. Reprinted in Fodor and Katz 1964, 334-352, Makkai 1972, 380-392.
- Halle, Morris. 1964. 'On the bases of phonology', in Fodor and Katz 1964, 324-333, Makkai 1972, 393-400.
- Halle, Morris. 1970. 'Theoretical Issues in phonology in the 1970's,' Proceedings of the Seventh International Congress of Phonetic Sciences, ed. by Andre Rigault and Rene Charbonneau.
- Halle, Morris. 1971. 'Is Kabardian a vowel-less language?,' Foundations of Language 6.1: 95-102.
- Halle, Morris and Stevens, Kenneth. 1971. 'A note on Laryngeal features', Quarterly Progress Report No. 101, Electronics Laboratory, M.I.T.
- Harms, Robert T. 1966. 'The measurement of phonological economy,' Language 42.III, 602-611.
- Hockett, Charles F. 1954. 'Two models of grammatical description' in Joos 1963, 386-399.
- Hockett, Charles F. 1955. A Manual of Phonology, IJAL, 21.4 (Pt.1).
- Hockett, Charles F. 1958. A Course in Modern Linguistics, New York, MacMillan.
- Hockett, Charles F. 1965. 'Sound change,' Language 41.2, 185-204.

- Householder, Fred W. 1965. 'On some recent claims in phonological theory', Journal of Linguistics 1.13-34.  
Also in Makkai 1972, 442-456.
- Hutchinson, Larry G. 1972. 'Mr. Chomsky on the phoneme,'  
reproduced by Indian University Linguistics Club.
- Hyman, Larry M. 1970. 'How concrete is phonology?,' Language  
46.I 58-76.
- Hyman, Larry M. 1975. Phonology: Theory and Practice,  
Holt, Rinehart and Winston.
- Jakobson, Roman. 1968. Child Language, Aphasia and Phonological  
Universals, (tr. by A. R. Keiler), The Hague.
- Johns, David, A. 1969. 'Phonemics and generative phonology,'  
Papers from the Fifth Regional Meeting of the Chicago  
Linguistic Society, 374-381. Also in Makkai 1972,  
549-553.
- Joos, Martin (ed.). 1963. Readings in Linguistics, American  
Council of Learned Societies, New York.
- Junghare, Indira Y. 1969. A Generative Approach to Marathi  
Vadhbhava Phonology, Unpublished M.A. dissertation,  
University of Texas, Austin.
- Kachru, Y. 1966. An Introduction to Hindi Syntax, Urbana,  
University of Illinois.
- Kelkar, Ashok R. 1968. Studies in Hindi-Urdu I: Introduction  
and word Phonology, Deccan College, Postgraduate and  
Research Institute, Poona.
- Kellogg, S.H. (Rev.). 1955 (1875). A Grammar of the Hindi  
Language, London, Routledge and Kegan Paul Ltd.
- King, Robert D. 1969. Historical Linguistics and Generative  
Grammar. Prentice-Hall Inc., Englewood Cliffs,  
New Jersey.
- Kiparsky, Paul. 1965. Phonological Change, Unpublished Ph.D.  
dissertation, M.I.T., reproduced by Indiana University  
Linguistics Club.
- Kiparsky, Paul. 1968a. 'Linguistic universals and linguistic  
change,' in Bach and Harms 1968, 170-202.
- Kiparsky, Paul. 1968b. 'How abstract is phonology?,'  
reproduced by Indiana University Linguistics Club,  
Bloomington, Indiana.

- Kiparsky, Paul. 1971. 'Historical linguistics', in Dingwall 1971, 576-649.
- Kiparsky, Paul. 1972. 'Explanation in Phonology,' in Peters 1972, 189-225.
- Kiparsky, Paul. 1973a. 'Abstractness, opacity, and global rules,' reproduced by Indiana University Linguistics Club.
- Kiparsky, Paul. 1973b. 'Elsewhere in phonology,' reproduced by the Indiana University Linguistic Club.
- Kisseberth, Charles W. 1969. 'On the abstractness of phonology: the evidence from Yawalmāni,' Papers in Linguistics 1.248-282.
- Kisseberth, Charles W. 1970a. 'On the functional unity of phonological rules,' Linguistic Inquiry, 1.291-306.
- Kisseberth, Charles W. 1970b. 'The treatment of exceptions', Papers in Linguistics 2.44-58.
- Kisseberth, Charles W. and Kenstowicz, Michael. 1973. 'Unmarked bleeding orders,' in Studies in Generative Phonology, 1-12, ed. by C. W. Kisseberth, Edmonton, Linguistic Research Inc.
- Koutsoudas, Andreas. 1973. 'Unordered rule hypothesis,' reproduced by the Indiana University Linguistic Club.
- Labov, William. 1969. 'Contraction, deletion, and inherent variability of the English copula,' Language 45.IV 715-762.
- Lass, Roger. 1971. 'Boundaries as obstruents: Old English voicing assimilation and universal strength hierarchies,' Journal of Linguistics 7.1:15-30.
- Lightner, Theodore M. 1971. 'Generative phonology' in Dingwall 1971, 499-574.
- Lockwood, David G. 1969. 'Markedness in stratificational phonology', Language 45.300-308. Also in Makkai 1972, 649-655.
- Maddah, Muhammad Mustafa Khan. 1972. Urdu-Hindi Shabda Kosh, Hindi Samiti, Lucknow.

- Makkai, Valerie B. (Ed.). 1972. Phonological theory: Evolution and Current Practice, Holt, Rinehart and Winston, Inc.
- Malone, Joseph L. 1970. 'In defense of non-uniqueness of phonological representations,' Language 46.II, 328-335.
- McCawley, James D. 1967. 'The role of a phonological feature system in a theory of language,' in Makkai 1972, 522-528.
- McGregor, R.S. 1968. The language of Indrajit of Orcha: A Study of Early Braj Bhasha Prose, Cambridge University Press.
- \* Misra, Bal Govind. 1967. Historical Phonology of Modern Standard Hindi: Proto-Indo-European to the Present, Unpublished Ph.D. dissertation, Cornell University.
- Narang, G.C., and Becker, D.A. 1971. 'Aspiration and nasalization in the generative phonology of Hindi-Urdu,' Language 47.III 646-667.
- \* Ohala, Manjari. 1972. Topics in Hindi-Urdu Phonology, Unpublished Ph.D. dissertation, UCLA.
- Ohala, Manjari. 1974. 'The abstractness controversy: experimental input from Hindi,' Language 50.II 225-235.
- Ohala, Manjari. 1974b. 'The Schwa-deletion rule in Hindi: Phonetic non Non-phonetic determinants of rule application,' reproduced by Indiana University Linguistics Club.
- Ohala, Manjari. 1975. 'Nasals and nasalization in Hindi,' Nasalfist: A Symposium on Nasals and Nasalization, 317-332, Stanford Language Universals Project.
- Ohala, M. and Ohala, J. 1972. 'The problem of aspiration in Hindi phonetics,' Annual Bulletin No. 6, Research Institute of Bogopedics and Phoniatics, University of Tokyo.
- Pandit, Prabodh B. 1961. 'Historical phonology of Gujarati vowels,' Language 37.54-66.
- Pandit, Prabodh B. 1976. 'A contribution to the schwa-deletion debate,' Indian Linguistics, 37.4, 282-295.

- Peters, Stanley (ed.). 1972. Goals of Linguistic Theory, Prentice-Hall.
- Pike, Kenneth L. 1947. 'Grammatical Prerequisites to Phonemic Analysis,' in Makkai 1972, 153-165.
- Pike, Kenneth L. 1952. 'More on Grammatical Prerequisites,' in Makkai 1972, 211-223.
- Postal, Paul M. 1968. Aspects of Phonological Theory, New York, Harper and Row.
- Postal, Paul M. 1970. 'On the surface verb 'remind',' Linguistic Inquiry, 1.37-120.
- Pray, Bruce R. 1970. Topics in Hindi-Urdu Grammar, Research Monograph No. 1, Center for South and Southeast Asia Studies, University of California, Berkeley.
- Sah, P.P. 1971. A Generative Semantic Treatment of Some Aspects of English and Hindi Grammar, Unpublished D.Phil. dissertation, University of York, England.
- Sapir, Edward. 1925. 'Sound patterns in language,' Language 1.37-51. Also in Makkai 1972, 13-21.
- Sapir, Edward. 1933. 'The Psychological Reality of Phonemes,' in Makkai 1972, 22-31.
- Saxena, Babu Ram. 1972. Avadhi ka Vikas, Hindustani Academy, Allahabad.
- Schane, Sanford A. 1968a. 'On the non-uniqueness of phonological representations,' Language 44.IV, 709-716.
- Schane, Sanford A. 1968b. French Phonology and Morphology, (Research Monograph, 45.) Cambridge, Mass., MIT Press.
- Schane, Sanford A. 1971. 'The phoneme revisited,' Language 47.III 503-21.
- Schane, Sanford A. 1972. 'Natural rules in phonology,' in Linguistic Change and Generative Theory, 199-229, ed. by R. P. Stockwell and K. S. MacCaulay, Bloomington, Indiana University Press.



- Schane, Sanford A. 1973. Generative Phonology, Prentice-Hall, Inc.
- Schane, Sanford A. 1974. 'How abstract is abstract?', in Papers from the Parasession on Natural Phonology (297-317), Chicago Linguistic Society.
- Sharma, Aryendra. 1975. A Basic Grammar of Modern Hindi, Central Hindi Directorate, Ministry of Education and Social Welfare, New Delhi.
- Shibatani, Masayoshi. 1971. 'The role of surface phonetic constraints in generative phonology,' POLA, Second Series, No. 13, MSI-MS54.
- Shukla, Shaligram. 1974. 'Phonological change and dialectal variation in middle Indo-Aryan,' in Historical Linguistics II, 391-402, ed. by J. M. Anderson and Charles Jones, North-Holland.
- Singh, Namvar. 1965. Hindi ke Vikas me Apabhramsa ka Yoga, Lok Bharti Prakashan, Allahabad.
- Sinha, A.C. 1970. Predicate Complement Constructions in English and Hindi, Unpublished D.Phil. dissertation, University of York, England.
- Sloat, C., Taylor, S.H., and Hoard, J.E. 1978. Introduction to Phonology, Prentice-Hall.
- Srivastava, R. N. 1969. Review of Kelkar's Studies in Hindi-Urdu I, Language, 45.IV 913-927.
- Stanley, Richard. 1967. 'Redundancy Rules in Phonology,' Language 43.393-436.
- Tewari, Udaya Narain. 1970. Hindi Bhasha ka Udgam aur Vikas, Bharti Bhandar, Leader Press, Allahabad.
- Trubetzkoy, N.S. 1968. Introduction to the Principles of Phonological Descriptions, (Tr. L. A. Murray, ed. H. Bluhme), The Hague.
- Vachek, Josef. 1970. 'Remarks on the sound pattern of English,' in Selected Writings in English and General Linguistics, 304-310, by Josef Vachek, The Hague.

- Vajpayee, Kishori Das. 1977. Hindi Sabdanushasan, Nagri Pracharni Sabha, Kashi.
- Varma, Dharendra. 1973. Hindi Bhasha ka Itihas, Hindustani Academy, Allahabad.
- Varma, Siddheshwar. 1961. Critical Studies in the Phonetic Observations of Indian Grammarians, First Indian edition, Munshi Ram Manohar Lal, Delhi.
- Verma, M.K. 1966. A Synchronic Comparative Study of the Structure of the Noun Phrase in English and Hindi, Ph.D. dissertation, Ann Arbor, University of Michigan.
- Vennemann, T. 1970. 'The German velar nasal: a case for abstract phonology,' Phonetica 22.65-81.
- Wang, William S-Y. 1968. 'Competing changes as a cause of residue,' Language 45.I 9-27.
- Wang, William S-Y. 1971. 'Approaches to English phonology,' POLA, Second Series, No. 13, WW1-WW41.
- Whitney, William Dwight. 1960. Sanskrit Grammar, Cambridge, Mass. Harvard University Press.
- Woolley, Dale E. 1968. 'Some observations on stridency in English,' in Papers from the Fourth Regional Meeting of Chicago Linguistic Society, 141-145.

We have not been able to consult the items marked with an asterisk.

#### ADDENDUM

- Ferguson, Charles A., and Chowdhury, M. 1960. 'The phonemes of Bengali,' Language 36. I 22-59.
- Kelkar, Ashok. R. 1958. The Phonology and Morphology of Marathi unpublished Ph.D. dissertation, Cornell University, Ithaca, N.Y.

Zvelebil, Kamil. 1970. Comparative Dravidian Phonology, The Hague, Mouton.